

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 03:43:12 : Search time 3799 Seconds
(without alignments)
9015.042 Million cell updates/sec

Title: US-10-715-129-1

Perfect score: 732

Sequence: 1 ttgtctcccccattcccca.....gaaattatcatatgatgat 732

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

EST: *
1: gb_ect1: *
2: gb_ect2: *
3: gb_ect3: *
4: gb_ect4: *
5: gb_ect5: *
6: gb_ect6: *
7: gb_ect7: *
8: gb_ect8: *
9: gb_ect9: *
10: gb_ect10: *
11: gb_ect11: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	531.4	72.6	543	1	A1998680
2	430	58.7	478	11	CR356994 Arabidops
3	401	54.8	477	9	BH753057
4	364.6	49.8	385	3	BP851800
5	262.4	35.8	920	10	CL485455
6	239	32.7	639	9	CC963171
7	191	26.1	321	9	BH904895
8	191	26.1	371	9	BH904894
9	182.6	24.9	471	1	AV544303
10	181	24.7	763	4	CNS0A8DR
11	140.8	19.2	616	9	BH528419
12	118	16.1	850	4	CNS0A7C
13	114.8	15.7	793	9	BZ03389
14	102.8	14.0	474	7	CV432287
15	99.6	13.6	765	9	CC737187
16	98	13.4	585	9	CC737307
17	96.4	13.2	663	9	CC737238
18	96.4	13.2	904	6	CA994079
19	93.2	12.7	646	6	CC737303
20	91.2	12.5	679	3	BH578473
21	88	12.0	621	9	CC737288
22	86.6	11.8	369	7	CNS48499

c 23	85.4	11.7	291	10	BX657886
c 24	85.2	11.6	622	7	CN893349
c 25	85	11.6	578	6	CR404276
c 26	83.2	11.4	595	9	CC737194
c 27	83.2	11.4	602	9	CC737269
c 28	82.8	11.3	382	2	BG125370
c 29	82.8	11.3	624	1	A1774560
c 30	82.6	11.3	606	3	B1934590
c 31	82.4	11.3	694	7	CV255174
c 32	82.4	11.3	771	7	CV26016
c 33	82.4	11.3	817	7	CV243132
c 34	80.8	11.0	758	8	CK655014
c 35	79.8	10.9	598	8	CK175107
c 36	79.8	10.9	635	7	CN193375
c 37	79.8	10.9	657	7	CN192376
c 38	79.8	10.9	661	7	CV476718
c 39	79.8	10.9	678	8	DN587342
c 40	79.8	10.9	749	8	DN625498
c 41	79.8	10.9	752	7	CF831325
c 42	79.8	10.9	775	7	CN187913
c 43	79.8	10.9	779	7	CN188780
c 44	79.8	10.9	859	8	CK045872
c 45	79.6	10.9	678	5	BW688619

ALIGNMENTS

RESULT 1
LOCUS A1998680/c 543 bp mRNA linear EST 08-SEP-1999
DEFINITION 701546552 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis
ACCESSION thaliana cDNA clone 701546552, mRNA sequence.
VERSION A1998680
KEYWORDS A1998680.1 GI:5845585

ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE
AUTHORS Chen J., Montoya M., Chan E., Mooney M., Carroon B., Gilliland D., Wang X., Hillman U., Giegler K., Kim C., Doyle M., Brzozka P., Gorgone G., Burns D., Grifflin J., Mounoudouva M., Nguyen D., Tan R., Rose M., Warren B., Ton B., Kaestury K., Borillo C., Carpio T., Policky J., Suzuki G., Argentine C., Shah S., Nobrega A., Murry L., Turner C., Krikorian S., Elder L. and Hanson D.

TITLE Arabidopsis thaliana Gene Expression Microarray
JOURNAL Unpublished (1999)

COMMENT Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: serv@genomesystems.com.
Location/Qualifiers

FEATURES

1. 543
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="701546552"
/issue_type="rosette"
/dev_str="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, rosette-2"
/note="Vector: pSPORT, Site 1: NotI, Site 2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and

watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunted, ligated to SalI adaptor, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN

Query Match 72.6%; Score 531.4; DB 1; Length 543;
Best Local Similarity 98.5%; Pred. No. 1,7e-115;
Matches 532; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 192 CTTACCAATGCTTCGATCAAGCTGATGTTAACAGTTGCCCTAAAGATCAGATCT 251
Db CTTACCAATGCTTCGATCAAGCTGATGTTAACAGTTGCCCTAAAGATCAGATCT 484
QY 252 TTGAGCAAAATTTTGTGATGATCTTTTAAATCCAAACGACACACCAAAACAACTCT 311
Db TTGAGCAAAATTTTGTGATGATCTTTTAAATCCAAACGACACACCAAAACAACTCT 424
QY 312 GTAGATGATTCGAGAAATCTCAAACTTACAAAAGATATTAAACATTCAGACAGTTA 371
Db GTAGATGATTCGAGAAATCTCAAACTTACAAAAGATATTAAACATTCAGACAGTTA 364
QY 372 TTGCAACACCGAGTCATGACGTCGAGAGAAACACCGAGAAACATGAGCTTTCGAG 431
Db TTGCAACACCGAGTCATGACGTCGAGAGAAACACCGAGAAACATGAGCTTTCGAG 304
QY 432 TTGCGCCGAGAAAAAGCAAGAGATTAACGCGGAGATTTTCGCGCGGAGAAAAAGAT 491
Db TTGCGCCGAGAAAAAGCAAGAGATTAACGCGGAGATTTTCGCGCGGAGAAAAAGAT 244
QY 492 GATCCCGGCAATATTATTAGTTAGAGTCTCTGTTCTATGTTGGTCTTAAAGGATC 551
Db GATCCCGGCAATATTATTAGTTAGAGTCTCTGTTCTATGTTGGTCTTAAAGGATC 184
QY 552 TCTGTTAAATCTTCGTTAGTTTTCGCGCGTTACCTCCGCTTCGTTTATGCTGAT 611
Db TCTGTTAAATCTTCGTTAGTTTTCGCGCGTTACCTCCGCTTCGTTTATGCTGAT 124
QY 612 GGTCCCATTTGGGATTAATGTTTACCTGCTGCTTCCTTCATGCTTCCTCATTC 671
Db GGTCCCATTTGGGATTAATGTTTACCTGCTGCTTCCTTCATGCTTCCTCATTC 64
QY 672 TAATGCTAATACAGATGATTAATCTGCAATTTCAATTAATCTGAATTTATTAATGA 731
Db TAATGCTAATACAGATGATTAATCTGCAATTTCAATTAATCTGAATTTATTAATGA 4

RESULT 2
CR356994/c 478 bp DNA linear GSS 05-APR-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-716H04-025111,
DEFINITION genomic survey sequence.
ACCESSION CR356994
VERSION CR356994.1 GI:45539916
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoever, P. and Weisshaar, B.
TITLE GABI-Kat Simplesearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana

JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
PUBMED 12874060
REFERENCE Rosso, M.G., Li, Y., Strizhov, N., Reis, B., Dekker, K. and Weisshaar, B.
AUTHORS An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics

JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
PUBMED 14756321
REFERENCE Strizhov, N., Li, Y., Rosso, M.G., Viehoever, P., Dekker, K.A. and Weisshaar, B.

JOURNAL High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
PUBMED Biotechniques 35 (6), 1164-1168 (2003)
REFERENCE 14682050

4 (bases 1 to 478)

REFERENCE Rosso, M.G., Strizhov, N., Li, Y. and Weisshaar, B.

JOURNAL Direct Submision
TITLE Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforshung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

COMMENT This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At3g59910. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genome program designated 'GABI'. Information on line availability can be found at: <http://www.mpi-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

1..478
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GK-716H04-025111"
/ecotype="Col-0"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (Ti) which were transformed with the T-DNA from vector pGABI (GenBank accession number: AY292716). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN

Query Match 58.7%; Score 430; DB 11; Length 478;

Best Local Similarity 100.0%; Pred. No. 1.7e-107;

Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 ACAACCTGTAGATGATTCGAGAAATCTCAAACTTACAAAAGATTTAATAACTTCA 362
Db ACAACCTGTAGATGATTCGAGAAATCTCAAACTTACAAAAGATTTAATAACTTCA 419
QY 363 AGACGTTATTGCAACACCGAGTCATGACGTCGAGAGAAACACCGAGAAACATGAG 422
Db AGACGTTATTGCAACACCGAGTCATGACGTCGAGAGAAACACCGAGAAACATGAG 359
QY 418 AGACGTTATTGCAACACCGAGTCATGACGTCGAGAGAAACACCGAGAAACATGAG 359
QY 423 CTTTCGAAATGTCGCGGAGAAAGCAAGAGATTAACGCGGAGTTTCTCGCGCGAGAA 482
Db CTTTCGAAATGTCGCGGAGAAAGCAAGAGATTAACGCGGAGTTTCTCGCGCGAGAA 299
QY 483 AAGAGATGATCCCGGAGATTAATTTCAATTTAGAGTCTGTTCTTATTTGTTGGTCT 542
Db AAGAGATGATCCCGGAGATTAATTTCAATTTAGAGTCTGTTCTTATTTGTTGGTCT 239
QY 298 AAGAGATGATCCCGGAGATTAATTTCAATTTAGAGTCTGTTCTTATTTGTTGGTCT 239
Db AAGAGATGATCCCGGAGATTAATTTCAATTTAGAGTCTGTTCTTATTTGTTGGTCT 179
QY 603 GCTGCTATTTGTTCCCATTTGGGATTAATGTTTACTGCTGTTCTTCAATGCTTTC 662
Db GCTGCTATTTGTTCCCATTTGGGATTAATGTTTACTGCTGTTCTTCAATGCTTTC 119
QY 178 GCTGCTATTTGTTCCCATTTGGGATTAATGTTTACTGCTGTTCTTCAATGCTTTC 119
Db GCTGCTATTTGTTCCCATTTGGGATTAATGTTTACTGCTGTTCTTCAATGCTTTC 722
QY 663 TTTCATTTCAATGCTAATACAGATGATTAATCTGCAATTTCAATTAATCTGAATTTAT 722
Db TTTCATTTCAATGCTAATACAGATGATTAATCTGCAATTTCAATTAATCTGAATTTAT 59
QY 723 ATATGATGAT 732
|||||||

Db 58 ATATGATGAT 49

RESULT 3
BH753057/c 477 bp DNA linear GSS 27-FEB-2002

LOCUS
DEFINITION
SALK_019707.56.00.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_019707.56.00.x, genomic
survey sequence.

ACCESSION
BH753057
VERSION
BH753057.1 GI:18972604

KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (chale crese)

DEFINITION
Arabidopsis thaliana (chale crese)

REFERENCE
AUTHORS
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R.,
Gadriab, C., Jeake, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
Shim, P., Zimmerman, J., and Ecker, J.R.

TITLE
A sequence-indexed library of insertion mutations in the
Arabidopsis Genome

JOURNAL
COMMENT
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

FEATURES
source
1..477
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_019707.56.00.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html

ORIGIN
Query Match 54.8%; Score 401; DB 9; Length 477;
Best Local Similarity 100.0%; Pred. No. 1.8e-99;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 CAAACTTCAAAAAGATTTTAAACATTCAAGACGTTATTCGAACACGAGTCATCG 391
Db 476 CAAACTTCAAAAAGATTTTAAACATTCAAGACGTTATTCGAACACGAGTCATCG 417
Qy 392 ACGTCGGAAGAAACACCGGAAAAACATGAGCTTTCGAAGTTCCGCGGAGAAAGCAAGC 451
Db 416 ACGTCGGAAGAAACACCGGAAAAACATGAGCTTTCGAAGTTCCGCGGAGAAAGCAAGC 357
Qy 452 AAGAGTTACGGCGAGTTTCGCGCGAGAAAAAGATGATCCCGCGAATTATTTCA 511
Db 356 AAGAGTTACGGCGAGTTTCGCGCGAGAAAAAGATGATCCCGCGAATTATTTCA 297
Qy 512 GTTTAGATCTCTGTTCTATTTGGTTGGTTCAACGGCATCTCTGTTAATTAATCTCGTTAG 571
Db 296 GTTTAGATCTCTGTTCTATTTGGTTGGTTCAACGGCATCTCTGTTAATTAATCTCGTTAG 237
Qy 572 TTTTGGCGCGGTTACCTCGGCTCGGTTATGCTGTTGGTTCCGATGGGAATTATAG 631

Db 236 TTTTGGCGCGGTTACCTCGGCTCGGTTATGCTGTTGGTTCCGATGGGAATTATAG 177

Qy 632 TTTTACTGCTGCTCTTCTGCTTCAATGCTTTCTCATTTCAATGCTAATACAGATGTA 691
Db 176 TTTTACTGCTGCTCTTCTGCTTCAATGCTTTCTCATTTCAATGCTAATACAGATGTA 117

Qy 692 CTTCGAATTCATGTAATATCGAAATTTATATATATATATATATATATATATATATAT 732
Db 116 CTTCGAATTCATGTAATATCGAAATTTATATATATATATATATATATATATATATAT 76

RESULT 4
BP851800
LOCUS
DEFINITION
BP851800 RAFI21 Arabidopsis thaliana cDNA clone RAFI25-17-007 5',
mRNA sequence.

ACCESSION
BP851800
VERSION
BP851800.1 GI:59912773

KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (chale crese)

DEFINITION
Arabidopsis thaliana (chale crese)

REFERENCE
AUTHORS
Seki, M., Ishida, J., Kamiya, A., Setou, M., Nakajima, M., Akiyama, K.,
Iida, K., Enju, A., Sakurai, T., Arahawa, T., Carninci, P., Fukuda, S.,
Iida, J., Kawai, J., Sasaki, D., Shiraki, T., Hayashizaki, Y., and
Shinozaki, K.

TITLE
Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs

JOURNAL
COMMENT
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekior@cc.riken.go.jp

FEATURES
source
1..385
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFI25-17-007"
/lab_host="DH10B"
/clone_lib="RAFI21"
/note="Site 1: BamHI; Site 2: SalI; Subtraction library.
The sequence was obtained from samples subjected to
various stresses and plant hormones-treated"

ORIGIN
Query Match 49.8%; Score 364.6; DB 3; Length 385;
Best Local Similarity 98.7%; Pred. No. 1.9e-89;
Matches 378; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 TCTTCTCATTTTCCCTACTAGTACTTGTGTTACACAGTTTCTTGTATCCAAACCAACCA 63
Db 1 TCTTCTCATTTTCCCTACTAGTACTTGTGTTACACAGTTTCTTGTATCCAAACCAACCA 60
Qy 64 TACACAAGCTTCAAACTCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 123
Db 61 TACACAAGCTTCAAACTCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 120
Qy 124 AGTTAACTCGAATTTGTTCTGATCTCTGTTTCTGAATGCTGGCCCAATCTTAATTTTGT 183
Db 121 AGTTAACTCGAATTTGTTCTGATCTCTGTTTCTGAATGCTGGCCCAT-CTTAATTTTGT 179
Qy 184 CTGAATTTCTTCAACCAATGCTTGCATCAAGCTGATTTGTTAACCAAGTTGCCCTAAAGA 243

```

Db      180 CTGGAATCTTCACCAATGCTTCGATCAAGTCAGTGGTTAACGAGTTCCTTAAGA 239
Qy      244 TCGATCTTTGAGCAAAATTTTGTCAGTATCTTCTTAATCCAAACGACGACGCAAAA 303
Db      240 TCGATCTTTGAGCAAAATTTTGTCAGTATCTTCTTAATCCAAACGACGACGCAAAA 299
Qy      304 CAACCTCTGATAGATTCGAGAAATCTCAAACTTAACAAAAGATTTATTAACATTCAA 363
Db      300 CAACCTCTGATAGATTCGAGAAATCTCAAACTTAACAAAAGATTTATTAACATTCAA 359
Qy      364 GACAGTTATTCGAACCAACCGAGT 386
Db      360 GACAGTTATTCGAACCAACCGAGT 382

RESULT 5
CL485455
LOCUS    SA1L.416.D03.v1 SA1L Collection Arabidopsis thaliana genomic clone
DEFINITION
ACCESSION CL485455
VERSION    CL485455.1 GI:45962135
KEYWORDS   GSS.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
REFERENCE  1 (bases 1 to 920)
AUTHORS    Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
            Dietrich,B., Ho,P., Baccaden,J., Ko,C., Clarke,J.D., Cotton,D.,
            Bullis,D., Snell,J., Miguel,T., Hutchinson,D., Kimerly,B.,
            Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
            A high-throughput Arabidopsis reverse genetics system
            Plant Cell 14 (12), 2985-2994 (2002)
12468722
COMMENT    Contact: Sessions A
            Applied Trait Genetics
            Syngenta Biotechnology Inc.
            3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
            Email: allen.sessions@syngenta.com
            ABRC Stock Number CS819225; T-DNA left border flanking sequences of
            Syngenta Arabidopsis Insertion Library (SAIL) lines are available
            through the Arabidopsis Biological Resource Center (ABRC).
            Sequences represent a pool of amplified genomic regions and not
            single contiguous sequences.
            Class: TDNA tagged.
            Location/Qualifiers
                1..920
                /organism="Arabidopsis thaliana"
                /mol_type="genomic DNA"
                /ecotype="Columbia"
                /db_xref="taxon:3702"
                /clone="SAIL.416.D03.v1"
                /clone_1db="SAIL Collection"
                /note="T-DNA left border sequences were isolated using a
                modified TAIL-PCR strategy"

ORIGIN
Query Match      35.8%; Score 262.4; DB 10; Length 920;
Best Local Similarity 70.5%; Pred. No. 4e-61;
Matches 437; Conservative 0; Mismatches 171; Indels 12; Gaps 6;

Qy      45 TTGATCCAAACCAACCAATACACAAAGCTTTCAACTCTTCACCTCAAGCTTCTTC 104
Db      250 TTGATCAAGCAACTCCGTCACAAAGAGCTTGGAATCTTCATGCCCCAGCTACGCA 309
Qy      105 CTTTACATCATGATCGTTGATTAAGTACCTGGAT-TTGTTTGCATCCCTGTTCTGATC 163
Db      310 CTTTACGCGGAAAACCTACCTACCATTTCTACGAGGCTGAGACCTTAAGGCACTCAATC 369
Qy      164 GTGGGCGCATCTTATTTGTCTGCAATCTTCAACCAATTCCTGCATCAAGCTGCATTGG 223

```

```

Db      370 GAGAGGCATCTTATTTAAACGCTCCAAATTAATACCAAAATAGCGGACTAAACCCACTTCAG 429
Qy      224 TTAAACGATGGCCCTTAAGATCAAGATCTTGAACAAAATTTGTCACTGATCTTTCAAT 283
Db      430 ATACCCGAG-----ACAGCTTATTAACCTGCACATAGGGAGAGTGTTCCTCCG 482
Qy      284 CCAACCGACACAGCAAAAACACCTCTAGATGANTT-CGAGAAATCTGAACCTTCAA 342
Db      483 TCAATCCACACACAAACCAATCCTCCAGTAGATGATTAACGAAATTTCTAAACTTACAA 542
Qy      343 AAAGATATTAATAAATTCAGACAGTATTCGAACCAACCGAGTCATGACGTCGAAAGA 402
Db      543 AAGATCTTATTAACATTCAGACAGTATTAACGAACCAACCGAGTCATGACGTCGAAAGA 602
Qy      403 AACACCGGAAAAACATGAGCTTTCCAGATTCGCGCGGAGAAAAGCAACGAAGTTACGG 462
Db      603 AACACCC-GTAAACACATGAGCTTTCCAGATTCGCGCGGAGAAAAGCAACGAAGTTACGG 661
Qy      463 CGGAGTTTCTCGGCGGAGAAAGATGATCCGCGCAATTAATTAATTCAGTTAGAGTCT 522
Db      662 CGGAGTTTCTCGGCGGCA-AAAAAGATGATATCCGCGCA-ATTTCAGATTAGAGTCT 719
Qy      523 CTGTTCTATTTGTTGTTCTTAACGCAATCTCTGTTAATACCTTCGTTAGTTTGCAGCG 582
Db      720 CTGTTCTATTCAGGGGGGCTTAACGCAATCTCTGTTAATACCTTCAGTTAGTTTGCAGCG 779
Qy      583 TTACCTTCGCTCTCGTTATGCTGCTATGTTGTTCCATTCGGAATATGTTTACTGTC 642
Db      780 TTACCTTACGCTCTCTTATGCTGCTAAGGTTCCACCTGGAGCTATGAGTAAACCCCG 839
Qy      643 GTTCTTGCTTCATGAGCTTC 662
Db      840 GGAATGGGCTTAAACCTTAC 859

RESULT 6
CC963171/c
LOCUS    BO1EM18F_BO_1.4_1.6_KB_nuc Brassica oleracea genomic clone
DEFINITION
ACCESSION CC963171
VERSION    CC963171.1 GI:33812283
KEYWORDS   GSS.
SOURCE     Brassica oleracea
ORGANISM   Brassica oleracea
REFERENCE  1 (bases 1 to 639)
AUTHORS    Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
            Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
            Whole genome shotgun sequencing of Brassica oleracea and its
            application to gene discovery and annotation in Arabidopsis
            Genome Res. 15 (4), 487-495 (2005)
15805490
COMMENT    Other_GSSs: BO1EM18F
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
            Location/Qualifiers
                1..639
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="TO1000DH3"
                /db_xref="taxon:3712"
                /clone="BO1EM18"

```

ORIGIN

/clone_11b="BO 1.4 1.6 KB nuc"
/note="Vector: pHO52; Site 1: BstXI; 1.4-1.6 kb sheared
nuclear DNA inserted into pHO52 using BstXI linkers"

Query Match 32.7%; Score 239; DB 9; Length 639;
Best Local Similarity 80.1%; Pred. No. 1.1e-54;
Matches 326; Conservative 0; Mismatches 60; Indels 21; Gaps 3;

ORIGIN

/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_105296.54.75.x"
/clone_11b="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html"

Query Match 26.1%; Score 191; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.7e-41;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

316 ATGATTGAGAAATCTCAACTTACAAAAGATTATTAACATTCAAGACATTATTG 375
317 |||||
318 |||||
319 |||||
320 |||||
321 |||||
322 |||||
323 |||||
324 |||||
325 |||||
326 |||||
327 |||||
328 |||||
329 |||||
330 |||||
331 |||||
332 |||||
333 |||||
334 |||||
335 |||||
336 |||||
337 |||||
338 |||||
339 |||||
340 |||||
341 |||||
342 |||||
343 |||||
344 |||||
345 |||||
346 |||||
347 |||||
348 |||||
349 |||||
350 |||||
351 |||||
352 |||||
353 |||||
354 |||||
355 |||||
356 |||||
357 |||||
358 |||||
359 |||||
360 |||||
361 |||||
362 |||||
363 |||||
364 |||||
365 |||||
366 |||||
367 |||||
368 |||||
369 |||||
370 |||||
371 |||||
372 |||||
373 |||||
374 |||||
375 |||||
376 |||||
377 |||||
378 |||||
379 |||||
380 |||||
381 |||||
382 |||||
383 |||||
384 |||||
385 |||||
386 |||||
387 |||||
388 |||||
389 |||||
390 |||||
391 |||||
392 |||||
393 |||||
394 |||||
395 |||||
396 |||||
397 |||||
398 |||||
399 |||||
400 |||||
401 |||||
402 |||||
403 |||||
404 |||||
405 |||||
406 |||||
407 |||||
408 |||||
409 |||||
410 |||||
411 |||||
412 |||||
413 |||||
414 |||||
415 |||||
416 |||||
417 |||||
418 |||||
419 |||||
420 |||||
421 |||||
422 |||||
423 |||||
424 |||||
425 |||||
426 |||||
427 |||||
428 |||||
429 |||||
430 |||||
431 |||||
432 |||||
433 |||||
434 |||||
435 |||||
436 |||||
437 |||||
438 |||||
439 |||||
440 |||||
441 |||||
442 |||||
443 |||||
444 |||||
445 |||||
446 |||||
447 |||||
448 |||||
449 |||||
450 |||||
451 |||||
452 |||||
453 |||||
454 |||||
455 |||||
456 |||||
457 |||||
458 |||||
459 |||||
460 |||||
461 |||||
462 |||||
463 |||||
464 |||||
465 |||||
466 |||||
467 |||||
468 |||||
469 |||||
470 |||||
471 |||||
472 |||||
473 |||||
474 |||||
475 |||||
476 |||||
477 |||||
478 |||||
479 |||||
480 |||||
481 |||||
482 |||||
483 |||||
484 |||||
485 |||||
486 |||||
487 |||||
488 |||||
489 |||||
490 |||||
491 |||||
492 |||||
493 |||||
494 |||||
495 |||||
496 |||||
497 |||||
498 |||||
499 |||||
500 |||||
501 |||||
502 |||||
503 |||||
504 |||||
505 |||||
506 |||||
507 |||||
508 |||||
509 |||||
510 |||||
511 |||||
512 |||||
513 |||||
514 |||||
515 |||||
516 |||||
517 |||||
518 |||||
519 |||||
520 |||||
521 |||||
522 |||||
523 |||||
524 |||||
525 |||||
526 |||||
527 |||||
528 |||||
529 |||||
530 |||||
531 |||||
532 |||||
533 |||||
534 |||||
535 |||||
536 |||||
537 |||||
538 |||||
539 |||||
540 |||||
541 |||||
542 |||||
543 |||||
544 |||||
545 |||||
546 |||||
547 |||||
548 |||||
549 |||||
550 |||||
551 |||||
552 |||||
553 |||||
554 |||||
555 |||||
556 |||||
557 |||||
558 |||||
559 |||||
560 |||||
561 |||||
562 |||||
563 |||||
564 |||||
565 |||||
566 |||||
567 |||||
568 |||||
569 |||||
570 |||||
571 |||||
572 |||||
573 |||||
574 |||||
575 |||||
576 |||||
577 |||||
578 |||||
579 |||||
580 |||||
581 |||||
582 |||||
583 |||||
584 |||||
585 |||||
586 |||||
587 |||||
588 |||||
589 |||||
590 |||||
591 |||||
592 |||||
593 |||||
594 |||||
595 |||||
596 |||||
597 |||||
598 |||||
599 |||||
600 |||||
601 |||||
602 |||||
603 |||||
604 |||||
605 |||||
606 |||||
607 |||||
608 |||||
609 |||||
610 |||||
611 |||||
612 |||||
613 |||||
614 |||||
615 |||||
616 |||||
617 |||||
618 |||||
619 |||||
620 |||||
621 |||||
622 |||||
623 |||||
624 |||||
625 |||||
626 |||||
627 |||||
628 |||||
629 |||||
630 |||||
631 |||||
632 |||||
633 |||||
634 |||||
635 |||||
636 |||||
637 |||||
638 |||||
639 |||||
640 |||||
641 |||||
642 |||||
643 |||||
644 |||||
645 |||||
646 |||||
647 |||||
648 |||||
649 |||||
650 |||||
651 |||||
652 |||||
653 |||||
654 |||||
655 |||||
656 |||||
657 |||||
658 |||||
659 |||||
660 |||||
661 |||||
662 |||||
663 |||||
664 |||||
665 |||||
666 |||||
667 |||||
668 |||||
669 |||||
670 |||||
671 |||||
672 |||||
673 |||||
674 |||||
675 |||||
676 |||||
677 |||||
678 |||||
679 |||||
680 |||||
681 |||||
682 |||||
683 |||||
684 |||||
685 |||||
686 |||||
687 |||||
688 |||||
689 |||||
690 |||||
691 |||||
692 |||||
693 |||||
694 |||||
695 |||||
696 |||||
697 |||||
698 |||||
699 |||||
700 |||||
701 |||||
702 |||||
703 |||||
704 |||||
705 |||||
706 |||||
707 |||||
708 |||||
709 |||||
710 |||||
711 |||||
712 |||||
713 |||||
714 |||||
715 |||||
716 |||||
717 |||||
718 |||||
719 |||||
720 |||||
721 |||||
722 |||||
723 |||||
724 |||||
725 |||||
726 |||||
727 |||||
728 |||||
729 |||||
730 |||||
731 |||||
732 |||||
733 |||||
734 |||||
735 |||||
736 |||||
737 |||||
738 |||||
739 |||||
740 |||||
741 |||||
742 |||||
743 |||||
744 |||||
745 |||||
746 |||||
747 |||||
748 |||||
749 |||||
750 |||||
751 |||||
752 |||||
753 |||||
754 |||||
755 |||||
756 |||||
757 |||||
758 |||||
759 |||||
760 |||||
761 |||||
762 |||||
763 |||||
764 |||||
765 |||||
766 |||||
767 |||||
768 |||||
769 |||||
770 |||||
771 |||||
772 |||||
773 |||||
774 |||||
775 |||||
776 |||||
777 |||||
778 |||||
779 |||||
780 |||||
781 |||||
782 |||||
783 |||||
784 |||||
785 |||||
786 |||||
787 |||||
788 |||||
789 |||||
790 |||||
791 |||||
792 |||||
793 |||||
794 |||||
795 |||||
796 |||||
797 |||||
798 |||||
799 |||||
800 |||||
801 |||||
802 |||||
803 |||||
804 |||||
805 |||||
806 |||||
807 |||||
808 |||||
809 |||||
810 |||||
811 |||||
812 |||||
813 |||||
814 |||||
815 |||||
816 |||||
817 |||||
818 |||||
819 |||||
820 |||||
821 |||||
822 |||||
823 |||||
824 |||||
825 |||||
826 |||||
827 |||||
828 |||||
829 |||||
830 |||||
831 |||||
832 |||||
833 |||||
834 |||||
835 |||||
836 |||||
837 |||||
838 |||||
839 |||||
840 |||||
841 |||||
842 |||||
843 |||||
844 |||||
845 |||||
846 |||||
847 |||||
848 |||||
849 |||||
850 |||||
851 |||||
852 |||||
853 |||||
854 |||||
855 |||||
856 |||||
857 |||||
858 |||||
859 |||||
860 |||||
861 |||||
862 |||||
863 |||||
864 |||||
865 |||||
866 |||||
867 |||||
868 |||||
869 |||||
870 |||||
871 |||||
872 |||||
873 |||||
874 |||||
875 |||||
876 |||||
877 |||||
878 |||||
879 |||||
880 |||||
881 |||||
882 |||||
883 |||||
884 |||||
885 |||||
886 |||||
887 |||||
888 |||||
889 |||||
890 |||||
891 |||||
892 |||||
893 |||||
894 |||||
895 |||||
896 |||||
897 |||||
898 |||||
899 |||||
900 |||||
901 |||||
902 |||||
903 |||||
904 |||||
905 |||||
906 |||||
907 |||||
908 |||||
909 |||||
910 |||||
911 |||||
912 |||||
913 |||||
914 |||||
915 |||||
916 |||||
917 |||||
918 |||||
919 |||||
920 |||||
921 |||||
922 |||||
923 |||||
924 |||||
925 |||||
926 |||||
927 |||||
928 |||||
929 |||||
930 |||||
931 |||||
932 |||||
933 |||||
934 |||||
935 |||||
936 |||||
937 |||||
938 |||||
939 |||||
940 |||||
941 |||||
942 |||||
943 |||||
944 |||||
945 |||||
946 |||||
947 |||||
948 |||||
949 |||||
950 |||||
951 |||||
952 |||||
953 |||||
954 |||||
955 |||||
956 |||||
957 |||||
958 |||||
959 |||||
960 |||||
961 |||||
962 |||||
963 |||||
964 |||||
965 |||||
966 |||||
967 |||||
968 |||||
969 |||||
970 |||||
971 |||||
972 |||||
973 |||||
974 |||||
975 |||||
976 |||||
977 |||||
978 |||||
979 |||||
980 |||||
981 |||||
982 |||||
983 |||||
984 |||||
985 |||||
986 |||||
987 |||||
988 |||||
989 |||||
990 |||||
991 |||||
992 |||||
993 |||||
994 |||||
995 |||||
996 |||||
997 |||||
998 |||||
999 |||||
1000 |||||
1001 |||||
1002 |||||
1003 |||||
1004 |||||
1005 |||||
1006 |||||
1007 |||||
1008 |||||
1009 |||||
1010 |||||
1011 |||||
1012 |||||
1013 |||||
1014 |||||
1015 |||||
1016 |||||
1017 |||||
1018 |||||
1019 |||||
1020 |||||
1021 |||||
1022 |||||
1023 |||||
1024 |||||
1025 |||||
1026 |||||
1027 |||||
1028 |||||
1029 |||||
1030 |||||
1031 |||||
1032 |||||
1033 |||||
1034 |||||
1035 |||||
1036 |||||
1037 |||||
1038 |||||
1039 |||||
1040 |||||
1041 |||||
1042 |||||
1043 |||||
1044 |||||
1045 |||||
1046 |||||
1047 |||||
1048 |||||
1049 |||||
1050 |||||
1051 |||||
1052 |||||
1053 |||||
1054 |||||
1055 |||||
1056 |||||
1057 |||||
1058 |||||
1059 |||||
1060 |||||
1061 |||||
1062 |||||
1063 |||||
1064 |||||
1065 |||||
1066 |||||
1067 |||||
1068 |||||
1069 |||||
1070 |||||
1071 |||||
1072 |||||
1073 |||||
1074 |||||
1075 |||||
1076 |||||
1077 |||||
1078 |||||
1079 |||||
1080 |||||
1081 |||||
1082 |||||
1083 |||||
1084 |||||
1085 |||||
1086 |||||
1087 |||||
1088 |||||
1089 |||||
1090 |||||
1091 |||||
1092 |||||
1093 |||||
1094 |||||
1095 |||||
1096 |||||
1097 |||||
1098 |||||
1099 |||||
1100 |||||
1101 |||||
1102 |||||
1103 |||||
1104 |||||
1105 |||||
1106 |||||
1107 |||||
1108 |||||
1109 |||||
1110 |||||
1111 |||||
1112 |||||
1113 |||||
1114 |||||
1115 |||||
1116 |||||
1117 |||||
1118 |||||
1119 |||||
1120 |||||
1121 |||||
1122 |||||
1123 |||||
1124 |||||
1125 |||||
1126 |||||
1127 |||||
1128 |||||
1129 |||||
1130 |||||
1131 |||||
1132 |||||
1133 |||||
1134 |||||
1135 |||||
1136 |||||
1137 |||||
1138 |||||
1139 |||||
1140 |||||
1141 |||||
1142 |||||
1143 |||||
1144 |||||
1145 |||||
1146 |||||
1147 |||||
1148 |||||
1149 |||||
1150 |||||
1151 |||||
1152 |||||
1153 |||||
1154 |||||
1155 |||||
1156 |||||
1157 |||||
1158 |||||
1159 |||||
1160 |||||
1161 |||||
1162 |||||
1163 |||||
1164 |||||
1165 |||||
1166 |||||
1167 |||||
1168 |||||
1169 |||||
1170 |||||
1171 |||||
1172 |||||
1173 |||||
1174 |||||
1175 |||||
1176 |||||
1177 |||||
1178 |||||
1179 |||||
1180 |||||
1181 |||||
1182 |||||
1183 |||||
1184 |||||
1185 |||||
1186 |||||
1187 |||||
1188 |||||
1189 |||||
1190 |||||
1191 |||||
1192 |||||
1193 |||||
1194 |||||
1195 |||||
1196 |||||
1197 |||||
1198 |||||
1199 |||||
1200 |||||
1201 |||||
1202 |||||
1203 |||||
1204 |||||
1205 |||||
1206 |||||
1207 |||||
1208 |||||
1209 |||||
1210 |||||
1211 |||||
1212 |||||
1213 |||||
1214 |||||
1215 |||||
1216 |||||
1217 |||||
1218 |||||
1219 |||||
1220 |||||
1221 |||||
1222 |||||
1223 |||||
1224 |||||
1225 |||||
1226 |||||
1227 |||||
1228 |||||
1229 |||||
1230 |||||
1231 |||||
1232 |||||
1233 |||||
1234 |||||
1235 |||||
1236 |||||
1237 |||||
1238 |||||
1239 |||||
1240 |||||
1241 |||||
1242 |||||
1243 |||||
1244 |||||
1245 |||||
1246 |||||
1247 |||||
1248 |||||
1249 |||||
1250 |||||
1251 |||||
1252 |||||
1253 |||||
1254 |||||
1255 |||||
1256 |||||
1257 |||||
1258 |||||
1259 |||||
1260 |||||
1261 |||||
1262 |||||
1263 |||||
1264 |||||
1265 |||||
1266 |||||
1267 |||||
1268 |||||
1269 |||||
1270 |||||
1271 |||||
1272 |||||
1273 |||||
1274 |||||
1275 |||||
1276 |||||
1277 |||||
1278 |||||
1279 |||||
1280 |||||
1281 |||||
1282 |||||
1283 |||||
1284 |||||
1285 |||||
1286 |||||
1287 |||||
1288 |||||
1289 |||||
1290 |||||
1291 |||||
1292 |||||
1293 |||||
1294 |||||
1295 |||||
1296 |||||
1297 |||||
1298 |||||
1299 |||||
1300 |||||
1301 |||||
1302 |||||
1303 |||||
1304 |||||
1305 |||||
1306 |||||
1307 |||||
1308 |||||
1309 |||||
1310 |||||
1311 |||||
1312 |||||
1313 |||||
1314 |||||
1315 |||||
1316 |||||
1317 |||||
1318 |||||
1319 |||||
1320 |||||
1321 |||||
1322 |||||
1323 |||||
1324 |||||
1325 |||||
1326 |||||
1327 |||||
1328 |||||
1329 |||||
1330 |||||
1331 |||||
1332 |||||
1333 |||||
1334 |||||
1335 |||||
1336 |||||
1337 |||||
1338 |||||
1339 |||||
1340 |||||
1341 |||||
1342 |||||
1343 |||||
1344 |||||
1345 |||||
1346 |||||
1347 |||||
1348 |||||
1349 |||||
1350 |||||
1351 |||||
1352 |||||
1353 |||||
1354 |||||
1355 |||||
1356 |||||
1357 |||||
1358 |||||
1359 |||||
1360 |||||
1361 |||||
1362 |||||
1363 |||||
1364 |||||
1365 |||||
1366 |||||
1367 |||||
1368 |||||
1369 |||||
1370 |||||
1371 |||||
1372 |||||
1373 |||||
1374 |||||
1375 |||||
1376 |||||
1377 |||||
1378 |||||
1379 |||||
1380 |||||
1381 |||||
1382 |||||
1383 |||||
1384 |||||
1385 |||||
1386 |||||
1387 |||||
1388 |||||
1389 |||||
1390 |||||
1391 |||||
1392 |||||
1393 |||||
1394 |||||
1395 |||||
1396 |||||
1397 |||||
1398 |||||
1399 |||||
1400 |||||
1401 |||||
1402 |||||
1403 |||||
1404 |||||
1405 |||||
1406 |||||
1407 |||||
1408 |||||
1409 |||||
1410 |||||
1411 |||||
1412 |||||
1413 |||||
1414 |||||
1415 |||||
1416 |||||
1417 |||||
1418 |||||
1419 |||||
1420 |||||
1421 |||||
1422 |||||
1423 |||||
1424 |||||
1425 |||||
1426 |||||
1427 |||||
1428 |||||
1429 |||||
1430 |||||
1431 |||||
1432 |||||
1433 |||||
1434 |||||
1435 |||||
1436 |||||
1437 |||||
1438 |||||
1439 |||||
1440 |||||
1441 |||||
1442 |||||
1443 |||||
1444 |||||
1445 |||||
1446 |||||
1447 |||||
1448 |||||
1449 |||||
1450 |||||
1451 |||||
1452 |||||
1453 |||||
1454 |||||
1455 |||||
1456 |||||
1457 |||||
1458 |||||
1459 |||||
1460 |||||
1461 |||||
1462 |||||
1463 |||||
1464 |||||
1465 |||||
1466 |||||
1467 |||||
1468 |||||
1469 |||||
1470 |||||
1471 |||||
1472 |||||
1473 |||||
1474 |||||
1475 |||||
1476 |||||
1477 |||||
1478 |||||
1479 |||||
1480 |||||
1481 |||||
1482 |||||
1483 |||||
1484 |||||
1485 |||||
1486 |||||
1487 |||||
1488 |||||
1489 |||||
1490 |||||
1491 |||||
1492 |||||
1493 |||||
1494 |||||
1495 |||||
1496 |||||
1497 |||||
1498 |||||
1499 |||||
1500 |||||
1501 |||||
1502 |||||
1503 |||||
1504 |||||
1505 |||||
1506 |||||
1507 |||||
1508 |||||
1509 |||||
1510 |||||
1511 |||||
1512 |||||
1513 |||||
1514 |||||
1515 |||||
1516 |||||
1517 |||||
1518 |||||
1519 |||||
1520 |||||
1521 |||||
1522 |||||
1523 |||||
1524 |||||
1525 |||||
1526 |||||
1527 |||||
1528 |||||
1529 |||||
1530 |||||
1531 |||||
1532 |||||
1533 |||||
1534 |||||
1535 |||||
1536 |||||
1537 |||||
1538 |||||
1539 |||||
1540 |||||
1541 |||||
1542 |||||
1543 |||||
1544 |||||
1545 |||||
1546 |||||

each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html

ORIGIN

Query Match 26.1%; Score 191; DB 9; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.7e-41;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGCTTCTCATTTCCCTACTAGTACTGTTTTCACACAGTTCTTGATCCAAACAAAC 60
DB 181 TTGCTTCTCATTTCCCTACTAGTACTGTTTTCACACAGTTCTTGATCCAAACAAAC 240
QY 61 CAATACACAAAGCTTCCAAAGCTTCCAAAGCTTCCAAAGCTTCCAAAGCTTCCAAAGCT 120
DB 241 CAATACACAAAGCTTCCAAAGCTTCCAAAGCTTCCAAAGCTTCCAAAGCTTCCAAAGCT 300
QY 121 TTGAGTTAAGTGGATTGTTGTCATCTCTGTTTGAATCGTGGCCATCTTATTT 180
DB 301 TTGAGTTAAGTGGATTGTTGTCATCTCTGTTTGAATCGTGGCCATCTTATTT 360
QY 181 TGTCTCGAATT 191
DB 361 TGTCTCGAATT 371

RESULT 9
AV544303/c 471 bp mRNA linear EST 20-FEB-2004
LOCUS AV544303 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION cDNA clone R239e09f 3', mRNA sequence.
ACCESSION AV544303
VERSION AV544303.1 GI:8715717
KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 471)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries

JOURNAL DNA Res. 7 (3), 175-180 (2000)
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 153-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES
SOURCE Location/Qualifiers
1..471
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="R239e09f"
/issue_type="roots"
/clone_id="Arabidopsis thaliana roots Columbia"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 24.9%; Score 182.6; DB 1; Length 471;
Best Local Similarity 70.0%; Pred. No. 3.8e-39;
Matches 296; Conservative 0; Mismatches 109; Indels 18; Gaps 3;

QY 270 CTGATCTTCTTAATCCAAACGACACGCAAAACCTCTGTAGATGATTCGAAAT 329
DB 460 CTGATCTTCTTAATCCAAACGACACGCAAAACCTCTGTAGATGATTCGAAAT 401

QY 330 CTCGAACCTTCAAAAAGATTATTAAACATTCAAGACAGTATTTCGAACACGAGCAT 389
DB 400 CTCGAACCTTCAAAAAGATTATTAAACATTCAAGACAGTATTTCGAACACGAGCAT 341
QY 390 GGAAGTCGGAAGA---AACACCGGAAGAAACATGAGCTTTCGAAGTT-----CGCC 437
DB 340 GGAAGTCGGAAGAAGATCATATACCGGAAGAAACAGAGTTTTCGAGTTTCAGCTCAGCTCC 281
QY 438 GGAAGAAAGACACGAAGATTACGGGAGGATTCTCGGCGCAGAA-----AGGATGAT 491
DB 280 GATTATGGGGAACCAAGAAATGTTTTCGACATGTCGTCCCAAGACAGTCCAGAGGCT 221
QY 492 GATCCGGGCAATTAATTTGATTAGAGTCTGTTCCATTGTTGGTGTGTAACGCGATC 551
DB 220 AATATACGAGGATTAATTCAGTTTGAATTAATGATGATGTTGTTGTTTCACAGATC 161
QY 552 TCTGTTAATACCTTCGTTAGTTTTCGCGGCTTACCTCGGCTTGAATGCTGATTT 611
DB 160 TCTGTTAATACCTTCGTTAGTTTTCGCGGCTTACCTCGGCTTGAATGCTGATTT 101
QY 612 GGTTCGATTTGGGATTAATGTTTACTGTCGTTCTTTCGTTCTATGCTTCTATTC 671
DB 100 GATTCTTATGGGATTAATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 41
QY 672 TAA 674
DB 40 CAA 38

RESULT 10
CNS0A8DR 763 bp mRNA linear HTC 06-FEB-2004
LOCUS CNS0A8DR
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSUTSIL73ZF02 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).

ACCESSION BX821785
VERSION BX821785.1 GI:42467792
KEYWORDS HTC; GSUT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 763)
Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Sallanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

JOURNAL Unpublished
PUBMED 2 (bases 1 to 763)
COMMENT Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage ;
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
Web : www.genoscope.cns.fr)

FEATURES
SOURCE The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Sallanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_FR/Full
length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.
Location/Qualifiers
1..763

ORIGIN
Query Match 24.9%; Score 182.6; DB 1; Length 471;
Best Local Similarity 70.0%; Pred. No. 3.8e-39;
Matches 296; Conservative 0; Mismatches 109; Indels 18; Gaps 3;

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTSL732F02"
/rname="Siliqua"
/ecotype="Col-0"
/plasmid="PCWSPORT_6"
1. .763
/gene="At2g44080"

ORIGIN

Query Match 24.7%; Score 181; DB 4; Length 763;
Best Local Similarity 69.7%; Pred. No. 1.2e-38;
Matches 295; Conservative 0; Mismatches 110; Indels 18; Gaps 3;

QY 270 CTGATCTTTCAATTCGAACGACGACGACAAACCACTGTGATGATTCGAAAT 329
DB 270 CTGATCTTTCAATTCGAACGACGACGACAAACCACTGTGATGATTCGAAAT 329
QY 330 CTGATCTTTCAATTCGAACGACGACGACAAACCACTGTGATGATTCGAAAT 389
DB 330 CTGATCTTTCAATTCGAACGACGACGACAAACCACTGTGATGATTCGAAAT 389
QY 390 GGAAGTCGAGAG--AACACCGGAAAAACATGAGCTTCGAGTT-----CGCC 437
DB 390 GGAAGTCGAGAGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 449
QY 438 GGAAGTCGAGAGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 491
DB 438 GGAAGTCGAGAGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 491
QY 450 GATATGAGGAGAGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 509
DB 450 GATATGAGGAGAGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 509
QY 492 GATCCCGGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 551
DB 492 GATCCCGGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 551
QY 510 AATATGAGGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 569
DB 510 AATATGAGGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 569
QY 552 TCTGTAATCTTCTGTTAGTTTGGCGCGCTACCTCCGCTCGTTTATGCTAT 611
DB 552 TCTGTAATCTTCTGTTAGTTTGGCGCGCTACCTCCGCTCGTTTATGCTAT 611
QY 570 TCTGTAATCTTCTGTTAGTTTGGCGCGCTACCTCCGCTCGTTTATGCTAT 629
DB 570 TCTGTAATCTTCTGTTAGTTTGGCGCGCTACCTCCGCTCGTTTATGCTAT 629
QY 612 GATTCGATGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 671
DB 612 GATTCGATGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 671
QY 630 GATTCGATGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 689
DB 630 GATTCGATGAGATTAACCGGAAAAACATGAGCTTCGAGTT-----CGCC 689
QY 672 TAA 674
DB 672 TAA 674
QY 690 CAA 692
DB 690 CAA 692

RESULT 11

BH528419 616 bp DNA linear GSS 13-DEC-2001
LOCUS BH528419.1
DEFINITION BOGC Braessica oleracea genomic clone BOGC46, genomic
survey sequence.
ACCESSION BH528419
VERSION BH528419.1 GI:17736504
KEYWORDS GSS.
SOURCE Braessica oleracea
ORGANISM Braessica oleracea
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 616)
Avele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,
Utecher, T.R., Mortman, J.R., White, O.R., and Town, C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)

JOURNAL
PUBMED
COMMENT

Other_GSSs: BOGC46TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-638-3523

Fax: 301-638-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. .616
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T0100DH3"
/db_xref="taxon:3712"
/clone="BOGC46"
/note="Vector: pHD51; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHD51 using BstXI linkers"

ORIGIN

Query Match 19.2%; Score 140.8; DB 9; Length 616;
Best Local Similarity 64.4%; Pred. No. 1.4e-27;
Matches 246; Conservative 0; Mismatches 127; Indels 9; Gaps 2;

QY 313 TAGATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 372
DB 313 TAGATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 372
QY 615 TAGATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 556
DB 615 TAGATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 556
QY 373 TCGAACCAAGATTCAGACAGTCGAGAAACACCGAAAAACATGAGCTTCGAGT 432
DB 373 TCGAACCAAGATTCAGACAGTCGAGAAACACCGAAAAACATGAGCTTCGAGT 432
QY 555 TCAAGCAGACCAACAGATTCAGACAGTCGAGAAAGATTAATTCGAGCTCAAGCT 496
DB 555 TCAAGCAGACCAACAGATTCAGACAGTCGAGAAAGATTAATTCGAGCTCAAGCT 496
QY 433 TCGCGGAGAAAGACAGAAAGATTAACCGGAGATTCGCGGAGAA-----AGG 486
DB 433 TCGCGGAGAAAGACAGAAAGATTAACCGGAGATTCGCGGAGAA-----AGG 486
QY 495 ---CTATGATGATTAACCAAAAGATTAATTCGAGCTTCGAGAAAGATTCGAGCTCAAG 439
DB 495 ---CTATGATGATTAACCAAAAGATTAATTCGAGCTTCGAGAAAGATTCGAGCTCAAG 439
QY 487 ATGATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 546
DB 487 ATGATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 546
QY 438 AGGATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 379
DB 438 AGGATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 379
QY 547 GCATCTCTGTAATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 606
DB 547 GCATCTCTGTAATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 606
QY 378 GCATCTCTGTAATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 319
DB 378 GCATCTCTGTAATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 319
QY 607 CTATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 666
DB 607 CTATGATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 666
QY 318 CTCTCAATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 259
DB 318 CTCTCAATTCGAGAAATTCGAACCAAAAGATTAATTAACATTCAGACAGTTAT 259
QY 667 CATTCTAATGCTAATCAGATG 688
DB 667 CATTCTAATGCTAATCAGATG 688
QY 258 TCCCTCTTATCTTCTAATG 237
DB 258 TCCCTCTTATCTTCTAATG 237

RESULT 12

CNS0AAT7C 850 bp mRNA linear HTC 06-FEB-2004
LOCUS CNS0AAT7C
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTFRH4ZB08 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX821044
VERSION BX821044.1 GI:42469690
KEYWORDS HTC; GSTR cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS
TITLE

Castelli, V., Anny, J.M., Jallion, O., Wincker, P., Clepet, C.,
Menard, M., Crnaud, C., Quetier, F., Scarpetti, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 850)

FEATURES	source
location/Qualifiers	1. 850
/organism="Arabidopsis thaliana"	
/mol_type="mRNA"	
/db_xref="taxon:3702"	
/clone="GSLTPG94ZE08"	
/tissue_type="Hormone Treated Callus"	
/ecotype="Col-0"	
/plasmid="PCWVS.P0RT_6"	
1. 850	
/gene="At2g44080"	

ORIGIN
gene

Query Match	16.1%	Score 118;	DB 4;	Length 850;
Best Local Similarity	61.7%	Pred. No. 3.1e-21;		
Matches 264;	Conservative 0;	Mismatches 145;	Indels 19;	Gaps 4;
270 CTGATCTCTTAATCAACGACAGACAGCAACCAACCTCGTATGATGATTCGGAAT	329			
377 CTGATTTTCAGATCAACCTTCTGACAGAGAGACCTCTTTAAATCATTCGTACAT	436			
330 CTCAAACTTACAAAAGATATTTATTAACATTCACAGATTAATTCGAAACAACGAGTAT	389			
437 CTCAGCTCTACAAAACGACATCATACATTCGAAAGAACATCATTCCTCCACAAACAT	496			
390 GGAGCTCGAAGAAAC--AACCGAAAACATGAGCTTTCGAGTTGCG-----CC 437				
497 GCACCTCGAGCCGATCATTAACCGACACACAGAGTTCGTGTTGAGCTCCAGCTCC	556			
438 GGAGAAAAGCAAGCAAGATTACGGCGAGTTCTCCGGCGAGAA-----AAGATGAT	491			
557 GATTATGGGAAAGCAAGAGTGTTCGACATTTGGTGGGCAAGAAAGTCCAAAGACCT	616			
492 GATCCCGCGAATTTATTTACGTTTGAAGTCTCT-GTTCCTAATGGTGTCTAACGAGAT	550			
617 CATATCAGACGAGTCACTCAGTTTGAATGAATGGTTTGCTGTGTGGGTCTGAAAAAT	676			
551 CTCTGTTAATCTTCGTTAGTTTTCGCGCGGTTACTTCGCGCTTCGTTATGCTCAT	610			
677 ATCTCTGATCTTACCGTTGATCTTCAACCAATGCTCTCTCTCTCTTTATGCTGATTT	736			
611 TGGTTCCGATTTGGGATTATGGTTTACCTGCTGCTCTTCCATATGCTCTCTCATTT	670			
737 TGATTTCCATTTGGGATTATGGTTTGTGCTTATGTTGGGCGCTGATGCGTGGGCGGATT	796			
671 CTAATGCT 678				
Db 797 TCAAGAT 804				

RESULT 13	LOCUS	DEFINITION
BZ039389	1k820f06.g1	B. oleracea002 Brassica oleracea genomic, genomic survey
BZ039389/c	793 bp	DNA linear GSS 09-OCT-2002

ACCESSION	BZ039389
VERSION	BZ039389.1
KEYWORDS	GSS.
SOURCE	Brassica oleracea
ORGANISM	Brassica oleracea
REFERENCE	Brassica oleracea
AUTHORS	Brassica oleracea
TITLE	Brassica oleracea
JOURNAL	Brassica oleracea
COMMENT	Brassica oleracea
FEATURES	Brassica oleracea
source	Brassica oleracea
Query Match	Brassica oleracea
Best Local Similarity	Brassica oleracea
Matches 149; Conservative	Brassica oleracea
Score 114.8; DB 9; Length 793;	Brassica oleracea
Pred. No. 2.3e-20;	Brassica oleracea
Mismatches 57; Indels 0; Gaps 0;	Brassica oleracea
480 GAAAGGATGATGATCCCGGAGTATTTTCAGTTTACGATCTCTGTTCCATTTGGTTGG	Brassica oleracea
513 GAGGAGGCTAATACTAATGAGTTAATCTTACGTTTGAAGTCAATGATGTTCTTGTGG	Brassica oleracea
540 TCTAACGCGATCTCTGTTAATCTCTCGTTAGTTTCCCGCGTTACCTCCGCTCCGT	Brassica oleracea
453 TCTACAGCATCGTCTTGTTCCTTCCTCGTTGATTTCTTCCACCGTGTCTCTCTCCCT	Brassica oleracea
600 TATGCTGATTTGTTCCATTTGGAATTTATGTTTATCTCGTGTCTTCTGCTTATGCG	Brassica oleracea
393 CATGCTGCTTCTGATTTCTTATGGAATTTATGTTTATGTTTCTTATGTTTCTTATGATG	Brassica oleracea
660 TTCTTCTATTTCTTATGTTTATGCG	Brassica oleracea
333 TTCTTCTATTTCTTATGTTTATGCG	Brassica oleracea
RESULT 14	Brassica oleracea
LOCUS	Brassica oleracea
DEFINITION	Brassica oleracea
ACCESSION	Brassica oleracea
VERSION	Brassica oleracea
KEYWORDS	Brassica oleracea
SOURCE	Brassica oleracea
ORGANISM	Brassica oleracea
REFERENCE	Brassica oleracea
AUTHORS	Brassica oleracea
TITLE	Brassica oleracea
JOURNAL	Brassica oleracea
COMMENT	Brassica oleracea

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 30, 2006, 08:59:59 ; Search time 3791 Seconds

(without alignments)
1308.212 Million cell updates/sec

Title: US-10-715-129-2

Perfect score: 533
Sequence: 1 MDVGRNKRKMSFRSSPEKS.....LAFMPSHSNANDVTGNCNM 106

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1p
-Q=/cgn2.1/USPRO.spool.p/US10715129/runat.27012006.154745.14579/app.query.fasta.1.263
-DB=EST -OPMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10715129 @CGN 1.1 8010 @runat.27012006.154745.14579 -NCPU=6 -ICPU=3
-NO MMAP -NBG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	533	100.0	477	9	BH753057 SALK_0197
2	533	100.0	478	11	CR356994 Arabidops
3	518	97.2	543	1	AI998680 701546552
4	395.5	74.2	639	9	CC963171 BOEIM187F
5	332	63.3	763	4	CNS0A8DR BX821785 Arabidops
6	327	61.4	471	1	AV544303 AV544303
7	292	54.8	616	9	BH528419 BOGCC46TR

c	8	274	51.4	793	9	BZ039389
c	9	270.5	50.8	382	2	BG125370 EST471016
c	10	270.5	50.8	624	1	AI774560 EST255660
c	11	259.5	48.7	678	5	BW688619
c	12	258.5	48.5	606	3	BI934590 EST554479
c	13	257.5	48.3	635	7	CN192375 UCRCS06
c	14	257.5	48.3	657	7	CN192376 UCRCS06
c	15	257.5	48.3	749	8	DN625498 UCRCA01
c	16	257.5	48.3	752	7	CP831325 UCRCS02
c	17	257.5	48.3	775	7	CN187913 UCRCS05
c	18	257.5	48.3	779	7	CN188780 UCRCS05
c	19	257.5	48.3	859	8	CX045872 UCRCS07
c	20	256.5	48.1	678	8	DN587342 UCRCS07
c	21	256.5	48.1	741	6	CP829407 UCRCS01
c	22	256	48.0	578	6	CP404276 CSBES024D
c	23	255.5	47.9	661	7	CV476718 25530.1 D
c	24	255.5	47.9	694	7	CV255174 WS02412.B
c	25	255.5	47.9	817	7	CV243132 WS02512.B
c	26	254.5	47.7	859	8	CX668291 UCRCP01
c	27	254.5	47.7	944	8	CX668915 UCRCP01
c	28	254.5	47.7	973	8	CX664941 UCRCP01
c	29	254	47.7	758	8	CX655014 P001003G0
c	30	254	47.7	771	7	CV266016 WS02028.B
c	31	253.5	47.6	369	7	CN548499 EST 16137
c	32	253.5	47.6	683	7	CF830612 UCRCS01
c	33	253.5	47.6	688	7	CF831943 UCRCS02
c	34	253.5	47.6	698	7	CF829409 UCRCS01
c	35	252.5	47.4	900	7	CF829408 UCRCS01
c	36	250.5	47.0	704	7	CF093413 G119P65.3
c	37	239.5	44.9	622	7	CF830612 UCRCS01
c	38	238.5	44.7	610	3	BP929940 BP929940
c	39	237.5	44.6	551	7	CV237738 WS0123.B2
c	40	237.5	44.6	598	8	CX175107 E10.69-17
c	41	237.5	44.6	706	8	CX174683 B02.69-19
c	42	237.5	44.6	794	7	CV257043 WS0245.B2
c	43	237	44.5	577	8	DN625497 UCRCA01
c	44	236.5	44.4	920	10	CI485455 SATL 416
c	45	234	43.9	581	2	BG317360 00487_1ea

ALIGNMENTS

RESULT 1
BH753057/c 477 bp DNA linear GSS 27-FEB-2002
LOCUS SALK_019707.56.00.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_019707.56.00.x, genomic
survey sequence.

ACCESSION BH753057
VERSION BH753057.1 GI:18972604
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadgilab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Becker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)

JOURNAL
COMMENT The Salk Institute Genomic Analysis Laboratory (SIGAL)
Contact: Joseph R. Becker
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel.: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckers@salik.edu

This is single pass sequence recovered from the left border of
TMDA. This sequence lies within 300 bases of the 3' end of
AT395900 and 300 bases of the 3' end of AT395910.

Class: TDNA tagged.

FEATURES
source

1..477
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_019707.56..00.x"
/clone_1db="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Alignment Scores:

Pred. No.: 8.17e-48 Length: 477
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-715-129-2 (1-106) x BH753057 (1-477)

Qy 1 MetAspValGlyArgGlnaAnaArgLysAanMetSerPheArgSerProGluLysSer 20
Db 420 ATGAGACGTGGGAAACAAACCGGAAACATGAGCTTTCGAAGTTCCGCGAGAAAAGC 361
Qy 21 LysGlnGluLeuArgArgSerPheSerAlaGlnLysArgMetCileProAlaAanTyr 40
Db 360 AAGCAAGATTACGCGCGAGTTTCTCGCGCGAGAAAGATGATGATCCGCGCAATTAT 301
Qy 41 PheSerLeuGluSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuLeuPro 60
Db 300 TTCAGTTTAAAGTCTCTGCTTCTATTGTTGTTAAACGCGATCTCTGTTAAATCTCCG 241
Qy 61 LeuValLeuProProLeuProProProPheMetLeuLeuLeuValProIleGlyIle 80
Db 240 TTACGTTTGGCGCGCTTACCTCCGCTCCGTTTATGCTGATGTTCCCATTTGGGATT 181
Qy 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAanAlaAanThrAsp 100
Db 180 ATGGTTTACTCGCGCTTCTTGGCTTATGCTTCTTCTATCTATATGTAATACGAT 121
Qy 101 ValThrCysAanPheMet 106
Db 120 GTAACCTGCATTTTCATG 103

RESULT 2

CR356994/c 478 bp DNA linear GSS 05-APR-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-716H04-025111,
DEFINITION genomic survey sequence.
CR356994
VERSION CR356994.1 GI:45539916
KEYWORDS GSS.

SOURCE
ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weishaar, B.
TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
the identification of T-DNA insertion mutants in Arabidopsis
thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)

JOURNAL
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and

TITLE Weishaar, B.
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics

JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
PUBMED 14756321
REFERENCE 3

AUTHORS Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and
Weishaar, B.

TITLE High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)

PUBMED 14682050

REFERENCE 4 (bases 1 to 478)

AUTHORS Rosso, M.G., Strizhov, N., Li, Y. and Weishaar, B.

TITLE Direct Submission

JOURNAL

COMMENT

Submitted (31-MAR-2004) Weishaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany.
This sequence has been recovered from the left border of the T-DNA.
It indicates an insertion close to or within gene AT3G59910.
Details on the protocols used for generation of the sequence are
described in References 1-3. The sequences are generated at the MPI
for Plant Breeding Research in the context of the GABI-Kat project.
GABI-Kat is part of the German Plant Genome program designated
'GABI'. Information on line availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

Location/Qualifiers

1..478

/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

/clone="GK-716H04-025111"

/clone_1db="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"

/note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector pGABI1 (Genbank accession number: AY529716). The
lines contain one or more T-DNA insertions. The DNA
fragment(s) resulting from the PCR were directly sequenced
to determine the genomic sequence flanking the insertion.
T-DNA derived sequences were removed."

ORIGIN

Alignment Scores:

Pred. No.: 8.19e-48 Length: 478
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-715-129-2 (1-106) x CR356994 (1-478)

Qy 1 MetAspValGlyArgGlnaAnaArgLysAanMetSerPheArgSerProGluLysSer 20
Db 393 ATGAGACGTGGGAAACAAACCGGAAACATGAGCTTTCGAAGTTCCGCGAGAAAAGC 334
Qy 21 LysGlnGluLeuArgArgSerPheSerAlaGlnLysArgMetCileProAlaAanTyr 40
Db 333 AAGCAAGATTACGCGCGAGTTTCTCGCGCGAGAAAGATGATGATCCGCGCAATTAT 274
Qy 41 PheSerLeuGluSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuLeuPro 60
Db 273 TTACGTTTAAAGTCTCTGCTTCTATTGTTGTTTAAACGCGATCTCTGTTAAATCTCCG 214
Qy 61 LeuValLeuProProLeuProProProPheMetLeuLeuLeuValProIleGlyIle 80
Db 213 TTACGTTTGGCGCGCTTACCTCCGCTCCGTTTATGCTGATGTTCCCATTTGGGATT 154
Qy 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAanAlaAanThrAsp 100
Db 153 ATGGTTTACTCGCGCTTCTTGGCTTATGCTTCTTCTATCTATATGTAATACGAT 94
Qy 101 ValThrCysAanPheMet 106

```

Db      93  GTAACTGCAATTCATG 76
|||||
RESULT 3
LOCUS   A1998680               543 bp  mRNA  linear  EST 08-SEP-1999
DEFINITION
thaliana cDNA clone 701546552, mRNA sequence.
ACCESSION
A1998680
VERSION
A1998680.1  GI:5845585
KEYWORDS
EST.
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 543)
AUTHORS
Chen, J., Momiyama, M., Chan, B., Mooney, M., Carroon, B., Gilliland, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzozka, P.,
Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D.,
Tan, R., Rose, M., Warren, B., Ton, B., Kasuri, K., Borrillo, C.,
Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,
Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and
Hanson, D.
Arabidopsis thaliana Gene Expression Microarray
Unpublished (1999)
CONTACT: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.
FEATURES
SOURCE
1. 543
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecoli_type="Col-0"
/db_xref="taxon:3702"
/clone="701546552"
/tissue_type="rosette"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, rosette-2"
/note="Vector: pSPORT, Site_1: NotI; Site_2: SalI; cDNA
library was derived from untreated rosette tissue from
Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks.
Plants were grown in 1:1:1 peat moss/vermiculite/perlite
soil at 22 deg. C +/- 3 deg. C under constant light, and
watered with fertilizer. cDNA synthesis was initiated
using a NotI-oligo(dT) primer. Double-stranded cDNA was
blunted, ligated to SalI adaptors, digested with NotI,
size-selected, and cloned into the NotI and SalI sites of
the pSPORT vector."
ORIGIN
Alignment Scores:
Pred. No.: 4,1e-46 Length: 543
Score: 518.00 Matches: 104
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 2
Query Match: 97.2% Indels: 0
Gaps: 0
DB: 1
US-10-715-129-2 (1-106) x A1998680 (1-543)
Qy      1  MetApYalGlyArgAnaAmrArgLyAsnMetSerPheArgSerSerProGlnLysSer 20
|||||
Db      347  ATGACCTCGGAAGAAACACCGGAAACATGACCTTCCAAATTTCCCGGAGAAAGC 288
|||||
Qy      21  LyGInGlnLueuArgArgSerPheSerAlaGlnLysArgMetCilieProAlaAsnTyr 40
|||||
Db      287  AAGCAAGATTACCGCGGAGATTCTCGCGCGCAAAAGATGATGATCCCGGCAATTAT 228
|||||

```

```

Qy      41  PheSerLengUisrLeuPheLeuValGlyLeuThrAlaSerLeuLeuLeuPro 60
|||||
Db      227  TTCAGTTAAGATCTCTGTTCTTAATGATGCTTAACGGCATCTCTGTAATCTCCG 168
|||||
Qy      61  LeuValLeuProProLeuProProProPheMetLeuLeuValProIleGlyIle 80
|||||
Db      167  TTACGTTTGGCGCGGCTTAACNTCCGCCCTCGTTATGCTGATTTGGTTCCCATTTGGAGTT 108
|||||
Qy      81  MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAlaAlaThrAsp 100
|||||
Db      107  ATGGTTTACTCGCGCTGCTTGGCTTCATGCTTCTCATCTTAATGCTTAATACAGAT 48
|||||
Qy      101  ValThrCysAsnPhenMet 106
|||||
Db      47  GTAACTGCAATTCATG 30
|||||
RESULT 4
LOCUS   CC963171/c              639 bp  DNA  linear  GSS 18-AUG-2003
DEFINITION
BOIEM18TF BO_1.4_1.6_KB_nuc Brassica oleracea genomic clone
BOIEM18, genomic survey sequence.
ACCESSION
CC963171
VERSION
CC963171.1  GI:33812283
KEYWORDS
GSS.
ORGANISM
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 639)
AUTHORS
Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,
Uteirback, T.R., Wortman, J.R., White, O.R. and Town, C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490
JOURNAL
PUBMED
COMMENT
Other_GSSs: BOIEM18TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Class: sheared ends.
FEATURES
SOURCE
1. 639
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOIEM18"
/clone_lib="BO_1.4_1.6_KB_nuc"
/note="Vector: pHO52; Site_1: BstXI; 1.4-1.6 kb sheared
nuclear DNA inserted into pHO52 using BstXI linkers"
ORIGIN
Alignment Scores:
Pred. No.: 1.05e-32 Length: 639
Score: 395.50 Matches: 82
Percent Similarity: 85.8% Conservative: 9
Best Local Similarity: 77.4% Mismatches: 10
Query Match: 74.2% Indels: 5
Gaps: 2
DB: 9
US-10-715-129-2 (1-106) x CC963171 (1-639)
Qy      4  GlyArgAnaAmrArgLyAsnMetSerPheArgSerSerPro-----GlnLys 19
|||||
Db      317  GGAAGAAACAACCGGAAACGTAACCTTTCGCGTCACCGGCAATGACGTAGAGAAC 258
|||||
Qy      20  SerLyGlnLueuArgArgSerPheSerAlaGlnLysArgMetCilieProAlaAsn 39
|||||

```


Pred. No.: 2.02e-25
 Score: 327.00
 Percent Similarity: 82.2%
 Best Local Similarity: 71.3%
 Query Match: 1
 DB: 1
 Length: 471
 Matches: 72
 Conservative: 11
 Mismatches: 12
 Indels: 6
 Gaps: 3

US-10-715-129-2 (1-106) x AV544303 (1-471)

Qy 1 MetaspVal---GlyAsgAsnAsnArgLysAsnMetSerPheArgSerSer-----16
 Db 342 ATGACGCGAGACGAGATCATACCGGAAACACGAGTTTTCGTTCAGCTCCAGCT 283
 Qy 17 ProGluLysSerLysGlnGluLeuArgSerPheSerAlaGlnLys-----ArgMet 34
 Db 282 CGGATTATGGGAGGAGCAAAATGTTTTCGACATTGTGTGCGAGAACAGTCCAGGAGG 223
 Qy 35 MetIleProAlaAsnTyrPheSerLeuGluSerLeuPheLeuValGlyLeuThrAla 54
 Db 222 CTAATATACGAGGAGTACTTCACTTACATCAATGATGCTTGTGCTTGTCTCAGCA 163
 Qy 55 SerLeuLeuIleLeuProLeuValLeuProLeuProProProPheMetLeuLeu 74
 Db 162 TCTCTCTGATCTTACCGTGTATCTTCCACCATGCTCCTCCTTATGCTGCTT 103
 Qy 75 LeuValProIleGlyIleMetValLeuLeuValValLeuAlaPheMetProSerSerHis 94
 Db 102 TTGATCTCATTTGGGATATATGTTTGTCTTATGTTCTTTCATGCTCTTATAT 43
 Qy 95 Ser 95
 Db 42 TCC 40

RESULT 7
 BMS28419/c 616 bp DNA linear GSS 13-DEC-2001
 LOCUS BMS28419
 DEFINITION BOCGJ46TR Brassica oleracea genomic clone BOCGJ46, genomic survey sequence.
 ACCESSION BMS28419
 VERSION BMS28419.1 GI:17736504
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 616)
 Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utecherback, T.R., Wortman, J.R., White, O.R. and Town, C.D.
 Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)
 15805490
 Other GSSs: BOCGJ46TR
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1. 616
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000D3"
 /db_xref="taxon:3712"
 /clone="BOGJ46"
 /note="Vector: pHOSt1. Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt1 using BstXI linkers"

ORIGIN

Alignment Scores:
 Pred. No.: 1.8e-21
 Score: 292.00
 Percent Similarity: 75.7%
 Best Local Similarity: 59.5%
 Query Match: 5
 DB: 5
 Length: 616
 Matches: 66
 Conservative: 18
 Mismatches: 19
 Indels: 8
 Gaps: 3

US-10-715-129-2 (1-106) x BMS28419 (1-616)

Qy 1 MetaspValGlyAsgAsnAsnArgLysAsnMetSerPheArgSerSerProGluLysSer 20
 Db 534 ATGACGCGAGACGAGATCATACCGGAAACACGAGTTTTCGTTCAGCTCCAGCT 484
 Qy 21 LysGlnGluLeuAlaGlySerPheSerAlaGlnLys-----ArgMetMetIleProAla 38
 Db 483 AAGCAAGATATTTTCCGACATTGTCTCAGAACAGTCCGAGAGGCTATATTCAGCG 424
 Qy 39 AsnTyrPheSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeu 58
 Db 423 AGTACTTCAAGTTTGGAGTCAATGCTTCTTGTGCTCAGAGCATCGCTTGATC 364
 Qy 59 LeuProLeuValLeuProProLeuProProProPheMetLeuLeuValProIle 78
 Db 363 CTTCCTGATCTTCCACCGTGTCTTCCCTCCTCCTGATGCTGCTTCTTATTCAT 304
 Qy 79 GlyIleMetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsn 98
 Db 303 CGGATATGCTTGTGCTTATGTTCTTCTTATGCTTCTTCTTCTTCTTCTTATCT 244
 Qy 99 ThrAsp-----ValThrCysAsnMet 106
 Db 243 TCTATGCCAACAATGTAACAGAACTATATG 211

RESULT 8
 BZ039389/c 793 bp DNA linear GSS 09-OCT-2002
 LOCUS BZ039389
 DEFINITION lka20f06.g1 B. oleracea002 Brassica oleracea genomic, genomic survey sequence.
 ACCESSION BZ039389
 VERSION BZ039389.1 GI:23626714
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 793)
 Delehanty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T., Nash, W., Rabinowicz, P.D. and Wilson, R.K.
 Whole genome shotgun reads from Brassica oleracea unpublished (2002)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@wustl.edu
 Plate: lka20 row: f column: 06
 Seq primer: -28RPOT reverse
 Class: shotgun
 High quality sequence start: 77
 High quality sequence stop: 511.
 Location/Qualifiers
 1. 793
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lid="B. oleracea002"
 /note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea TO1000D3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic

FEATURES

source
 1. 616
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000D3"
 /db_xref="taxon:3712"
 /clone="BOGJ46"
 /note="Vector: pHOSt1. Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt1 using BstXI linkers"

DNA was provided by Pablo Rabnowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Alignment Scores: 2.22e-19 Length: 793
Pred. No.:

Score:	274.00	Matches:	64
Percent Similarity:	77.6%	Conservative:	19
Best Local Similarity:	59.8%	Mismatches:	17
Query Match:	51.4%	Indels:	2
DB:	9	Gaps:	2

US-10-715-129-2 (1-106) x BZ039389 (1-793

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

5 ArgasbnaabnargLybAbnmetSerpne

US-10-715-129-2 (1-106) x BZ039389 (1-793

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84


```

/clone lib="tomato resistant, Cornell"
/note="Vector: plbluescript SK(-); Site_1: EcoRI, Site_2:
XhoI; c1er - Tomato Pseudomonas Resistant EST library.
Directionally cloned cDNAs inserted into plbluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

```

Alignment Scores:

Pred. No.:	3,97e-19	Length:	62
Score:	270.50	Matches:	56
Percent Similarity:	77.9%	Conservative:	18
Best Local Similarity:	58.9%	Mismatches:	16
Query Match:	50.8%	Indels:	5
DB:	1	Gaps:	2

US-10-715-129-2 (1-106) x AI774560 (1-624)

Oy	13	PheATGSeSerProGluuyluSerlyrGNGluLeuAlaArgerPheSer---AlaGln	31
Db	332	TTCAACACATTAATGAGGAACATAGATGAGCATTAAAGATCATTTACACAGACAT	351
Oy	32	LybATgMeMetileProAlaenlyrPheSerLeuGluserLeuPheLeuValGly	51
Db	392	GCGCAAGAAATCTTGTCATATGATTTTTCAGCTTAGACATATTTTGTACTGTGT	451
Oy	52	LeuThraIaSerLeuLeuileuProLeuValLeuProLeuProProProPhe	71
Db	452	CTTAACGATCTTTGTACTTTTGCCATATGATGCTTCACACATGGCACCAACCTTT	511
Oy	72	MetLeuLeuLeuValProileGlyileMetValLeuLeuValLeuAlaPheMetPro	91
Db	512	ATGTTGTTGTAGTCCCATATTTTCATCTTGTTGTTCTTATGATCTTAGCTTTATGCT	571
Oy	92	SeSerIrlaSerAlaAlaSnhrAepValIthrCyAsnPhemet	106
Db	572	TCATAATGTTAGGAAT-----GGACATGTGCTCANATCTT	604

RESULT 11	
LOCUS	BM688619
DEFINITION	BM688619 678 bp mRNA linear EST 27-APR-2005 BM688619 Lycopodium esculentum cv. Micro-Tom fruit Lycopodium esculentum cDNA clone FC12CF10 5', mRNA sequence.

SOURCE ORGANISM	
lycopersicon esculentum	(Solanum lycopersicum)
lycopersicon esculentum	

REFERENCE
AUTHORS

Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 678)

Tsugane, T., Watanabe, M., Yano, K., Suzuki, H., Sakurai, N. and
Shibata, D.

ORIGIN

Alignment Scores:	
Pred. No.:	6, 9e-18
Score:	259, 50
Percent Similarity:	76.8%
Best Local Similarity:	55.8%
Query Match:	48.7%
DB:	5
length:	677
Matches:	53
Conservative:	20
Mismatches:	17
Indels:	1
Gaps:	2

US-10-715-129-2 (1-106) X BW688619 (1-678)

```

QY      13 PheArgSerSerProGluLysSerIysGlnGluLeuAlaArgSerPheSer---AlaGln 3
Db      359 TTCAACAACCTTATGCAAGGCAATAAATGAGCATATAAAATCATTTACACAGACACT 418
QY      32 LysArgMetMetLeuProAlaAsnTyrPheSerLeuGluSerLeuPheLeuValGly 51
Db      419 GGCAGAAAAAGTTGTCATGATGATTAATTTTACCTTAAGTCAATTAATTTGGTACTTGGT 478
QY      52 LeuThrAlaSerLeuLeuIleLeuProLeuValLeuProProLeuProProPhe 71
Db      479 CTTACAGCAATCTTTTACTTTTGCCATGTAAGCTTCCACCAATGGCACACCACTTTT 538
QY      72 MetLeuLeuLeuValProIleGlyLeuMetValLeuLeuValValLeuAlaPheMetPro 91
Db      539 ATGTTGTGTTAGCCCCCAATTTTCATCTCTGTGTTCTTATGATCTTACCTTTATGGCT 598
QY      92 SerSerHisSerAsnAlaAsnThrAspValThrCysAsnPheMet 106
Db      599 TCTAATGTTAGCAAT-----GAGACTTGCTCAATATCTT 631

```

RESULT 12				
BI934590/c				
LOCUS	BI934590	606 bp	mRNA	linear
DEFINITION	EST554479 tomato flower, anthesis Lycopersicon esculentum cDNA			
	clone cTOD2086 5' end, mRNA sequence.			

SOURCE ORGANISM	
<i>Lycopersicon esculentum</i>	(<i>Solanum lycopersicum</i>)
<i>Lycopersicon esculentum</i>	

REFERENCE
1 (bases 1 to 606)

FEATURES	Location/Qualifiers
source	1. .606

They were immediately frozen in liquid nitrogen and then

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosids; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 657)
Close, T.D., Roose, M.L., Arpaia, M.L., Federici, C.F., Fenton, R.D., Wamamaker, S., Focht, E., Sievert, J., Robinson, P., Kim, H.R., Kudrna, D., Stum, D., Yost, D. and Wing, R.
Development of EST Resources and New Genetic Markers for California Citrus - Washington Navel Orange Stored Fruit Rind

TITLE

Unpublished (2004)

JOURNAL

Contact: Timothy Close
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124
Tel: 9097873318
Fax: 9097874437
Email: timothy.close@ucr.edu
Seq primer: T3.

FEATURES

SOURCE

Location/Qualifiers
1..657
/organism="Citrus sinensis"
/mol_type="mRNA"
/cultivar="Washington navel"
/db_xref="taxon:2711"
/clone="CS_MEC006M17"
/tissue_type="Rind"
/dev_stage="Commercially producing trees"
/lab_host="E. coli TUC121"
/clone_id="Washington Navel Orange Stored Fruit Rind cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Fruits were collected
January-March 2003 [Federici, Roose lab, Focht, Sievert &
Robinson, Arpaia lab]. Four samples related to storage
conditions were produced: 1) Fresh-picked in Mentone
(Arnott Brothers Enterprises, Mentone, CA), 2) after 21
days storage at 5C at Kearney then transported to UC
Riverside on ice, 3) after 5 additional days storage at
11C at Kearney, sampled immediately, 4) fruit grown in
southern CA were obtained from Redlands foothill Packing
House after commercial packing, X-ray irradiated at 300 Gy
by Surebeam, then stored 1 day at ambient temperature.
Rind tissue (juice vesicles) were collected. Tissues were
snap frozen and then stored at -80C until further
processing. Fenton (Close lab) purified RNA by the phenol
method described in J. Japanese Soc. Hort. Sci. 1996. 64
(4): 809-814, purified poly(A) mRNA using a PolyAtract
mRNA Isolation System IV (Promega), produced a primary
cDNA library using a lambda ZAP XR cDNA Synthesis Kit
(Stratagene), then mass-excised one million pfu from the
primary library to produce a phagemid population.
Phagemids were plated, plasmid DNA purified, cDNA clones
archived, and DNA sequences determined bi-directionally
using an ABI3730 at the Arizona Genomics Institute,
University of Arizona (Kim, Kudrna, Stum, Yost, Wing).
Chromatogram files were downloaded by FTP to UC Riverside
(by Close), then processed at UC Riverside (by Wamamaker,
Close lab) using the HarvEST pipeline
(http://harvest.ucr.edu) to remove vector and cloning
oligo sequences and various contaminants, and to trim to a
high quality region. Sequences that retained a phred 17
region of at least 100 bases were deposited to GenBank."

ORIGIN

Alignment Scores:

Pred. No.: 1..1e-17 Length: 657
Score: 257.50 Matches: 56
Percent Similarity: 75.6% Conservative: 12
Best Local Similarity: 62.2% Mismatches: 17
Query Match: 48.3% Indels: 5
DB: 7 Gaps: 2

US-10-715-129-2 (1-106) x CN192376 (1-657)

Qy 18 GltysserlysgingluleuArgSerPheSer---AlaGlnlyArgMetMetIle 36
Db 47 GAGAGAGAGAAAGTTGATGATACCGAGATTACTCGCAAGGCGCATCCAGAAACGTGTT 106
Qy 37 ProHlaanTYrPheSerLeuGlusierleuPheleuValglyleuThrAlaSerleu 56
Db 107 TCGGCAAGCATATTCACCGCTGAAATCATTTGCTTTGTCGTATGCTACGCGCCATG 166
Qy 57 LeuileuProleuValleuProProleuProProleuProPheMetleuVal 76
Db 167 CTGATCCCTGCATTCGCTCCGCCCTTCGCCGCCGCCCATTCCTGCTCTGCTT 226
Qy 77 ProileglyIleMetValleuValValleuValPheMetProSerSerHisSerAsn 96
Db 227 CTTATGATATTCATGCGGCTTCCTTGTGCTTCGATTCATGCTCTTAAT----- 277
Qy 97 AlaasnThrAspValThrCysAsnPheMet 106
Db 278 ---GTAAGAGTATTAATCTTCACGTACGTG 304

RESULT 15
DN625498 749 bp. mRNA linear EST 12-APR-2005
LOCUS UCRCA01.06113 x Bark of Madame Vinous Sweet Orange Scion on
DEFINITION Standard Sour Orange Rootstock, Citrus Tristeza Virus Challenged -
UCRCA01 Citrus aurantium cDNA clone CA_SBA006113, mRNA sequence.
ACCESSION DN625498
VERSION DN625498.1 GI:61694598
KEYWORDS EST.
SOURCE Citrus aurantium
ORGANISM Citrus aurantium

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
Rosids; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 749)
Close, T.D., Roose, M.L., Federici, C.F., Mandala, J., Fenton, R.D.,
Kudrna, D., Wallen, C., Patino, A., Wamamaker, S., Kim, H.R.,
Muller, C., Collura, K., Misosetski, M., Byrne, M., Stum, D., Smart, D.,
Development of EST Resources and New Genetic Markers for California
Citrus - Bark of Madame Vinous Sweet Orange Scion on Standard Sour
Orange Rootstock, Challenged with Citrus Tristeza Virus cDNA
library UCRCA01

TITLE

JOURNAL

Unpublished (2005)
Contact: Timothy J. Close
Department of Botany & Plant Sciences
University of California
Riverside, CA 92521-0124, USA
Tel: 909-787-3318
Fax: 909-787-4437
Email: timothy.close@ucr.edu
Seq primer: T3.

FEATURES

SOURCE

Location/Qualifiers
1..749
/organism="Citrus aurantium"
/mol_type="mRNA"
/cultivar="Standard Sour Orange"
/db_xref="taxon:43166"
/clone="CA_SBA006113"
/tissue_type="bark below bud union, predominantly C.
aurantium"
/dev_stage="young trees"
/lab_host="E. coli TUC121"
/clone_id="Bark of Madame Vinous Sweet Orange Scion on
Standard Sour Orange Rootstock, Citrus Tristeza Virus
Challenged - UCRCA01"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; All plants were grown and
inoculated at the Central California Tristeza Broadcra-
tion Agency (CCTRA), under the direction of Marylou Polek, with
assistance by Cindy Wallen, who grew and inoculated the
plants, and Tony Patino who collected tissue and ran the
ELISA tests. All plants were Madame Vinous sweet orange

(*Citrus sinensis* Osbeck) budded onto standard Sour Orange rootstock (*Citrus aurantium* L.). They were inoculated onto the scion with bark patches infected with various strains of CTV. Two inoculation dates: 1/24/03, 7/14/04. Three collection dates: 8/19/04 (2 sets collected on this date), 10/21/04, and 12/9/04. Four plants 'days after inoculation', groups: 36, 99, 160, 573 days after inoculation. Number of plants collected for each set: 36-day (18 plants), 99-day (30 plants), 160-day (38 plants), and 573-day (18 plants). These included three virus-free control plants for each group, and 3 or more plants for each virus strain. Five quick-decline virus strains of CTV, identified by CTRBA isolate numbers, were used for the two sets of plants collected on 8/19/04: 107, 142, 96142, 96180, and 96208. Nine quick-decline virus strains of CTV were used for 10/21/04 and 12/9/04 collections: 107, 142, 96142, 96180, 96208, 160, 177, 146, 151. All of the 36-day plants failed to show positive ELISA tests of the scion before the sampling date, and continued to be negative in the rootstock when tested one and two months after the sampling date. The PCR test of the pooled RNA extracted from these plants was negative, so we had no convincing evidence that the virus was active at the bud union in these samples. Most of the 573 days-after-inoculation plants, which were growing outdoors, had tested ELISA positive for CTV at some time, but some of them gave negative tests shortly before sampling. However, they were showing decline symptoms, so all were judged to be infected. RNA was extracted from a pool of all the samples. This pool gave a strong RT-PCR positive result. ELISA tests indicated positive results for scion tissue from many of the plants of the 99-day and 160-day samples, but many were consistently negative. Some of the plants with negative ELISA scores did show some apparent symptoms of stunting, or stem pitting. RT-PCR was done by Claire Federici on some of the samples collected for the 99-day and 160-day samples, RNA from scion bark directly above bud-union sample piece. Those that were PCR tested had elevated ELISA scores, not quite positive, or they had symptoms that suggested they might be showing signs of CTV even though the ELISA filter did not indicate presence of the virus. Of the 27 99-day plants inoculated with CTV, 14 had positive ELISA tests. Of the 13 ELISA negative inoculated plants, 9 were tested by RT-PCR and 2 of these gave a positive result; the others were all negative by both types of test. Of the 35 160-day plants inoculated with CTV, 6 had positive ELISA tests. Of the 29 ELISA negative inoculated plants, 14 were tested by RT-PCR and all were negative. Which plants were used for the pools? All inoculated plants collected from the 36-day set and all inoculated plants from the 573-day set were used in two separate pools, regardless of ELISA scores. For the 99-day and 160-day plants, only ones with positive ELISA and/or RT-PCR results were used, and they were divided into four pools, with either high or low ELISA filter for each date. Samples were collected by Claire Federici, Mary Lou Polek, and Cindy Wallen. For each sample, a razor blade was used and the material collected in RNAlater (Ambion). The bud union with rootstock bark included the bark from just above the bud-union (this was the bottom cut of the scion bark piece) to about 1 cm below the bud union. That means this tissue included both rootstock bark and some scion bark. For every tree there were two tubes of bud union/rootstock. One bud union/rootstock tube of each tree was to be pooled for the cDNA library production. The other was to be saved for gene expression studies. There was one departure from this plan: the pools for the 160-day plants were made from both tubes of rootstock/bud union bark from these plants since the tissue amount was small. Samples in RNAlater were transported on ice to UC Riverside and then stored at -20 (until extracted. CTRBA contact information: Dr. Marylou Polek, Associate Plant Pathologist, Program Manager,

ccsee@lightspeed.net, Department of Food and Agriculture, Plant Health & Pest Prevention Services, Integrated Pest Control Branch, Central California Tribes Radiation Agency, 22847 Road 140, Tulare, CA 93274. Jayci Mandal and Raymond Fenton (Close lab) purified RNA by a TRIzol method, purified poly(A) mRNA using PolyAT Tract RNA Isolation Kit (Promega), produced a primary cDNA library using a lambda ZAP XR cDNA Synthesis Kit (Stratagene), then mass-excised 1 million pfu from the primary library to produce a phagemid population. The library was made from a mixture of equal quantities of RNA from each of the four sample types. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the Arizona Genomics Institute (Kim, Kidrne, Collura, Wiscocki, Byrne, Stum, Smart, Muller, Wing). Chromatogram files were downloaded by FTP by Close, then processed by Manamaker (Close lab) using the Harvest pipeline (<http://harvest.ucr.edu>) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were assembled, then chimeras were removed following manual inspection of assemblies (Manamaker, Close, Roose). Sequences that survived all removal steps were submitted to GenBank. Clones from this library are archived at the Arizona Genomics Institute (<http://www.genome.arizona.edu/orders/>).

ORIGIN

Alignment Scores:

Pred. No.: 1,29e-17 Length: 749
Score: 257.50 Matches: 56
Percent Similarity: 75.6% Conservative: 12
Best Local Similarity: 62.2% Mismatches: 17
Query Match: 48.3% Indels: 5
DB: 8 Gaps: 2

US-10-715-129-2 (1-106) x DN625498 (1-749)

```

QY      18 GILUYSerLyGInGUleuAgaRgSerPhSer---AlaGlnYsaRgMcMeIle 36
Db      135 GAGAAAGAGGAAAGATTGATTCACCGAAGTTACTCCGCAAGGGGCACTCAAAAAGTGT 194
QY      37 ProAlaantYrPhSerLeuGInuSerLeuPhaLeuLeuValGlyLeuThraIaSerLeu 56
Db      195 TCGGCAAGCAATTTCACCTCGAATCATGCTTTGCTCTATGCTTGAAGGCTCATTTG 254
QY      57 LeuIleLeuProLeuValLeuProProLeuProProProPhaMetLeuLeuVal 76
Db      255 CTGATCTGTCATTTGATGGCTTCGCCCTTGCGCCGCCGCCCATTCGCGCTTGCTT 314
QY      77 ProIleGlyIleMetValLeuLeuValValLeuAlaIlePhaMetProSerRhaSerAn 96
Db      315 CCTATGAGTATTTCTACCGCGCTTTGCTTGTGGCAATTCATGCGCTTAAAT----- 365
QY      97 AlaantThraPyAlaThrCysaanPhMet 106
Db      366 ---GTAAAGATATTAATCTTCACGATGATG 392

```

Search completed: January 30, 2006, 11:17:32
Job time : 3796 secs

ORIGIN

/translation="MDVGRNNRKNMFRSSPERSKDELRSFSQKRMIPANYSLE
SLFLVGLTASLILPLVPLPLPPPMALLVPGLVWLVLVAFMPSHNNATDVLT
CNFM"

Query Match 100.0%; Score 732; DB 15; Length 732;
Best Local Similarity 100.0%; Pred. No. 2,6e-176;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGCTCCATCTCCCTAGTACTGTTTCAACAGTTTCTGATCCAAACAAAC 60
DB 1 TTGCTCCATCTCCCTAGTACTGTTTCAACAGTTTCTGATCCAAACAAAC 60
QY 61 CAATACCAAAAGCTTCTCAAACTCTTCACTCAAGCTTCTTCACTCAATCG 120
DB 61 CAATACCAAAAGCTTCTCAAACTCTTCACTCAAGCTTCTTCACTCAATCG 120
QY 121 TTGAGTTAAGCTGGATTTGTTTGCATCTCTGTTTCTGAATCGTGGCCATCTTAATT 180
DB 121 TTGAGTTAAGCTGGATTTGTTTGCATCTCTGTTTCTGAATCGTGGCCATCTTAATT 180
QY 181 TTGTCGAATCTTCAACAAATGCTTGATCAAGCTGATGTTAACAAGTCCCTAA 240
DB 181 TTGTCGAATCTTCAACAAATGCTTGATCAAGCTGATGTTAACAAGTCCCTAA 240
QY 241 AGATCAGATCTTTGAGCAAAATTTTGTCACTGATCTTCAAACTCAACAGACAGCA 300
DB 241 AGATCAGATCTTTGAGCAAAATTTTGTCACTGATCTTCAAACTCAACAGACAGCA 300
QY 301 AAACAACTCTGATGATGATTCGAAATCTCAAACTTACAAAAGATATTATAACATT 360
DB 301 AAACAACTCTGATGATGATTCGAAATCTCAAACTTACAAAAGATATTATAACATT 360
QY 361 CAAGACAGTTATTCGAAACCGAGTCATGACGTCGGAAGAAACACCGGAAAAATG 420
DB 361 CAAGACAGTTATTCGAAACCGAGTCATGACGTCGGAAGAAACACCGGAAAAATG 420
QY 421 AGCTTTCGAAGTTGCGCGAGAAAGCAAGCAAGATTAACGCGGAGTTTCTCGCGCAG 480
DB 421 AGCTTTCGAAGTTGCGCGAGAAAGCAAGCAAGATTAACGCGGAGTTTCTCGCGCAG 480
QY 481 AAAAGATGATGATCCCGGCAATTAATTCAGTTAAGTCTGTTCTTAATGTTGGT 540
DB 481 AAAAGATGATGATCCCGGCAATTAATTCAGTTAAGTCTGTTCTTAATGTTGGT 540
QY 541 CTAAGGAGATCTGTTAATCTCCGTTAGTTTTCGCGCGTAACTCCGCGCTCGCTT 600
DB 541 CTAAGGAGATCTGTTAATCTCCGTTAGTTTTCGCGCGTAACTCCGCGCTCGCTT 600
QY 601 ATGCTGATATGTTGCCATTTGGAATTATGTTTACCTGCTGCTTCTTCAATGCT 660
DB 601 ATGCTGATATGTTGCCATTTGGAATTATGTTTACCTGCTGCTTCTTCAATGCT 660
QY 661 TCTTCTCATTTAATGCTAATACAGATGTAACCTTCAATTCATGTAATCTGAATTTA 720
DB 661 TCTTCTCATTTAATGCTAATACAGATGTAACCTTCAATTCATGTAATCTGAATTTA 720
QY 721 TTAATGATGAT 732
DB 721 TTAATGATGAT 732

RESULT 2
ATP24G16 100906 bp DNA linear PLN 16-Apr-2005
LOCUS ATP24G16 100906 bp DNA linear PLN 16-Apr-2005
DEFINITION Arabidopsis thaliana DNA chromosome 3, BAC clone F24G16.
ACCESSION AL138647
VERSION AL138647.1 GI:6899904
KEYWORDS
SOURCE Arabidopsis thaliana (chale crees)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

REFERENCE

1 D'Angelo, M., Vezzi, A., Modesto, D., Pigazzi, M., Valle, G.,
Nemes, H.M., Lemcke, K., Mayer, K.F.X., Quetier, F. and Salanoubat, M.
Unpublished
2 (bases 1 to 100906)
EU Arabidopsis sequencing project.
Direct Submission
Submitted (18-FEB-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
d'Interet Public, Centre National de Sequencage - GENOSCOPS; 2 rue
Gaston Cremieux, BP191, 91006 Evry Cedex, France;
http://www.genoscope.cns.fr

COMMENT

Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

FEATURES

source
1..100906
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/ecotype="Columbia"
1..1731
/note="overlap to BAC T16124, please refer EMBL Acc. no
AL138659 for analysis and annotation"
2..1888
/gene="F24G16.10"
/complement (<2..1888)
/gene="F24G16.10"
complement (<2..1888)
/gene="F24G16.10"
/note="Contains protein kinases signatures and profile
AA339-362; Protein kinases signatures and profile
AA458-470"
/product="receptor lectin kinase 3"
/protein_id="CA875793.1"
/db_xref="GI:7019668"
/db_xref="GOA:Q9M200"
/db_xref="InterPro:IPR000719"
/db_xref="InterPro:IPR000985"
/db_xref="InterPro:IPR001220"
/db_xref="InterPro:IPR008271"
/db_xref="UniProt/TrEMBL:Q9M200"
/translation="MSHKVQLQIVLVLLTFSSTHNSNGNFMERAAAAGLNGCYLLT
NTTKHSYGOAFNNTPVPKIKSFRSFPNIIRGIVPEHKQGSHGMAFYSPTRGIPGASP
DOYLGFNENNNKASNNVYAIELDIRKDEFGDIDNHGIVINGLTSVASAGAY
DDEIDGPKKLSISTVMRLSIYSHTDQNLNTLIPARISVPKSLISLNDLSAPY
FLAEYTLGTAISGSLGALTYNQSFEYSBYIIPANDLVITPLPPYPKSYRTARI
LAVCLTLAVFTALVAGSIGFVFRKRVKVELEWEIQLGPPRPYKELFNATKQPK
EKQLDLKRGQGVYKMLPGSDAEIIVKRTSHSRQSMBSFLAEITIGLRHPNIVR
LIGYCHKENLIVYDFMPNGSLDRCLTSNTSNEBERLTWSERFKIKDVAATALHL
HGWVGVIVHRDIKPAWVLLDHGMNRLIDFGIAKLYDGFQDTSRVTGATYIAP
LIRTERATSTDYVARGIYMEFYVCGRPRIIRERAAENAVLQMIIELEWESGLTPAA
EESIRQNRGRLEIVLKLGLCAHHTELIRPMASVLTQLNSVSHLPNNLIDVPAE
RLKGI"
complement (2..1888)
/gene="F24G16.10"
/number=1
2685..4651
/gene="F24G16.20"
complement (join(2685..4127,4214..4651))
/gene="F24G16.20"
complement (join(2685..4127,4214..4651))
/gene="F24G16.20"
/note="strong similarity to receptor lectin kinase 3 -
Arabidopsis thaliana, EMBL:U93161
Contains protein kinases signatures and profile
AA309-332; Protein kinases signatures and profile
AA425-437"

```

/codon_start=1
/product="receptor lectin kinase-like protein"
/protein_id="CAB75794.1"
/db_xref="GI:7019669"
/db_xref="GOA:Q9M129"
/db_xref="InterPro:IPR000719"
/db_xref="InterPro:IPR000985"
/db_xref="InterPro:IPR001220"
/db_xref="InterPro:IPR008271"
/db_xref="UniProt/TrEMBL:Q9M129"
/translation="MPSLKVLAHIVLVLVLTLLSSSTYNSNGMTLEGSAA DNSIGDTI
LVTNKKHSCQTFNNBSIPKIDSFSFHLFGIYPHTQSGHMSFVISEPTAGLPAA
SSQYGLFENSTGKSNHYAIELDIQKQEGDIDNHNVAWMLSIYSHPDOO
LNTLTPABIPVPRKPLSLNRDLSPLYLEMYGYTATGSI GAFHYMLSSVATRK
VENPTFRIVPTLPYPKSSDRTKILANCLLAFPAVAVAGICPVPTTRKTKK
EVLSEMTIQPHRFAYKELNATKDKREKOLKGGGQGVFKGTLPSNDELAVKT
SHSQRQMSSEFLAISTIGRLRPNLVLLGYCHKNLVLVDPFTNGSLDKYLDN
ENORLTWEORFKIKDVASALHQBWOI IHRDIKPAVLI DHEMNAIIGDFGL
AKLYDQGLDPTSRVAGTFGYIAPELRTGATSTVVAAGLVMLVVGCRMIEER
APREERYLVMIIEIMESGKLFDAAESIRQEQNRGRILILIKGLCAHHTELIRN
MSAVMOILNGVSQLPNDLDVVRANKRMPEISIEVLGLNLTYSVGTMTITNSFLSH
GR"
complement (2685..4127)
/gene="P24G16.20"
/number=1
complement (4128..4213)
/gene="P24G16.20"
/number=1
complement (4214..4651)
/gene="P24G16.20"
/number=2
5498..8175
/gene="P24G16.30"
complement (join(5498..5626,5840..5920,6042..6101,
6253..6332,6432..6483,6611..6748,7011..7274,7405..7462,
7549..7658,7819..8175))
/gene="P24G16.30"
complement (join(5498..5626,5840..5920,6042..6101,
6253..6332,6432..6483,6611..6748,7011..7274,7405..7462,
7549..7658,7819..8175))
/gene="P24G16.30"
/notes="several differences to EMBL:X81973
Contains ATP/GTP-binding site motif A (P-loop)
A169-176/Cysteine synthase/cystathionine beta-synthase
P-phosphate attachment site A143-161"
/codon_start=1
/product="cysteine synthase"
/protein_id="CAB75795.1"
/db_xref="GI:7019670"
/db_xref="GOA:Q43725"
/db_xref="UniProt/Swiss-Prot:Q43725"
/translation="MVAMIMASRPNRAKASQILSTLGNRSCTSMMAATSSALL
NPLTSSSSSTLRFRFCSPETISLSPSASDFSLAMRQSRSPADGSEDPSPVCEAV
KRETGPGLNIADNVSQLIGKTPMVYLNLSAKGCVANIAKLBIIMBCSCVNDIGYS
MTVDAEOKGFISPKSVLAVPTSGNTGIGLAFIASRGYLLITMPASMSMERVLK
ARGBAYLTPRAKGTCAVQKABEILKTKTPAYMLQDPNDPANFKIHRETTGPRIWD
TKGVLDLFAVGIGTGSTTGVGRFKKRNPTQVIGVEPESDILSGKQKPHIKQGI
GAGETPNLDQKMDVIAISSEBAISTAKOLAKELMGVSISSGAANAIRKAKRP
ENAGETLAVTVNBSMAVFOVFPFSFERYSTPLFQSIREEVERKQPEV"
complement (5498..5626)
/gene="P24G16.30"
/number=1
complement (5627..5839)
/gene="P24G16.30"
/number=1
complement (5840..5920)
/gene="P24G16.30"
/number=2
complement (5921..6041)
/gene="P24G16.30"
/number=2
complement (6042..6101)
/gene="P24G16.30"

```

Query Match	100.0%; Score 732; DB 15; Length 100906;
Best Local Similarity 100.0%; Pred. No. 1,9e-176;	
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 TTGTCCTCAATTCCTCACTAGTACTGTTTTCACACAGTTTCTTGATCCAAAC 60	/number=3 complement (6102..6252) /gene="P24G16.30"
DB 62556 TTGTCCTCAATTCCTCACTAGTACTGTTTTCACACAGTTTCTTGATCCAAAC 62615	/number=3 complement (6253..6332) /gene="P24G16.30"
QY 61 CAATACCAAAAGCTTCAAACTCCTCACTCAAGCTTCTTCTTAACTCGAATCG 120	/number=4 complement (6333..6431) /gene="P24G16.30"
DB 62616 CAATACCAAAAGCTTCAAACTCCTCACTCAAGCTTCTTCTTAACTCGAATCG 62675	/number=4 complement (6432..6483) /gene="P24G16.30"
QY 121 TTGAGTTAACTGGATTGTTTCTCATCTCTGTTTCTGAATCGTGAGCCATCCTTAATT 180	/number=5 complement (6484..6610) /gene="P24G16.30"
DB 62676 TTGAGTTAACTGGATTGTTTCTCATCTCTGTTTCTGAATCGTGAGCCATCCTTAATT 62735	/number=5 complement (6611..6748) /gene="P24G16.30"
QY 181 TGTCGAATTCCTCAACCAATTCCTCACTCAAGCTCAATGTTAAACCAATTGCCCTAA 240	/number=6 complement (6749..7010) /gene="P24G16.30"
DB 62736 TGTCGAATTCCTCAACCAATTCCTCACTCAAGCTCAATGTTAAACCAATTGCCCTAA 62795	/number=6 complement (7011..7274) /gene="P24G16.30"
QY 241 AGATCAATCTTGAAGAAAATTTTGCACTGATCTTCTTAAATCCAAACGACACGCA 300	/number=7 complement (7275..7404) /gene="P24G16.30"
DB 62796 AGATCAATCTTGAAGAAAATTTTGCACTGATCTTCTTAAATCCAAACGACACGCA 62855	/number=7 complement (7405..7462) /gene="P24G16.30"
QY 301 AAACAACCTCTGATGATGATTCGAGAATCTCAACTTAACAAAAGATATTATAACATT 360	/number=8 complement (7463..7548) /gene="P24G16.30"
DB 62856 AAACAACCTCTGATGATGATTCGAGAATCTCAACTTAACAAAAGATATTATAACATT 62915	/number=8 complement (7549..7658) /gene="P24G16.30"
	/number=9 complement (7659..7818) /gene="P24G16.30"
	/number=9 complement (7819..8175) /gene="P24G16.30"
	/number=10 complement (8176..12111) /gene="P24G16.40"

QY 361 CAAGACAGTTATTCGAACAAACCGAGTCATGAGCTCCGAAAGAAACACCGGAAAAACATG 420
Db 62916 CAAGACAGTTATTCGAACAAACCGAGTCATGAGCTCCGAAAGAAACACCGGAAAAACATG 62975
QY 421 AGCTTTGAAAGTTGCGCGAGAAAAGCAAGACAGATTACCGCGAGATTCTCGCGCAG 480
Db 62976 AGCTTTGAAAGTTGCGCGAGAAAAGCAAGACAGATTACCGCGAGATTCTCGCGCAG 63035
QY 481 AAAAGAGTAGATCCCGGGAATTAATTCAGTTAGAGTCTCGTCTCAATGGTTGGT 540
Db 63036 AAAAGAGTAGATCCCGGGAATTAATTCAGTTAGAGTCTCGTCTCAATGGTTGGT 63095
QY 541 CTACGCGCATCTCTGTATTAATCTCCGTAGTTTGGCGCGTACCTCCGCGCTCCGTT 600
Db 63096 CTACGCGCATCTCTGTATTAATCTCCGTAGTTTGGCGCGTACCTCCGCGCTCCGTT 63155
QY 601 ATGCTGTAATGTTGCCCATTTGGGATTAAGTTTACTCGTCTGTTGCTTCGCTTCATGCT 660
Db 63156 ATGCTGTAATGTTGCCCATTTGGGATTAAGTTTACTCGTCTGTTGCTTCGCTTCATGCT 63215
QY 661 TCTTCATTCATATGCTAATACAGATGTAATCTGCAATTCATGTAATCTGAATTTA 720
Db 63216 TCTTCATTCATATGCTAATACAGATGTAATCTGCAATTCATGTAATCTGAATTTA 63275
QY 721 TTATATGATGAT 732
Db 63276 TTATATGATGAT 63287

RESULT 3
AKI76640
LOCUS
DEFINITION Arabidopsis thaliana 747 bp mRNA linear PLN 09-SEP-2004
clone: RAFL25-17-J07.

ACCESSION AKI76640.1 GI:51971476
VERSION AKI76640.1
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,
Narusaka, M., Shin, I., Nakagawa, M., Sakamoto, N., Oishi, K.,
Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,
Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs
Unpublished
2 (bases 1 to 747)

TITLE
JOURNAL Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720).
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified lambda phage vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This
clone is in a modified phage vector.
Please visit our web site (http://range.gsc.riken.jp/) for further
details.

COMMENT
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720).
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified lambda phage vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This
clone is in a modified phage vector.
Please visit our web site (http://range.gsc.riken.jp/) for further
details.

FEATURES
SOURCE 1. 747
Location/Qualifiers

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAFL25-17-J07"
/ecotype="Columbia"
/note="common name: thale cress"
1. 747
/gene="At3g59900"
314. 706
/gene="At3g59900"
/codon_start=1
/product="putative protein"
/protein_id="BAD4403.1"
/db_xref="GI:51971477"
/translation="MIRKISNUKQDINIDSYNNRVNDVGNRRKMSFRSSPKS
KQELRFSFAQKMKIPANFSDISLPLVGLTASLLIPLVPLPPPLLVPI
GIMVLIVLAFPSHSHNANTVTCNFM"

Query Match 99.7%; Score 730; DB 15; Length 747;
Best Local Similarity 100.0%; Pred. No. 8,3e-176;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GTCCTCCTCATTTCCCTACTAGTACTGTTTTCACAGATTCTTGATCCACCAAAACCA 62
Db 1 GTCCTCCTCATTTCCCTACTAGTACTGTTTTCACAGATTCTTGATCCACCAAAACCA 60
QY 63 ATACACAAAGCTTTCGAACTCCTTCGCAAGCTTCTCTTTACATCGAATCGTT 122
Db 61 ATACACAAAGCTTTCGAACTCCTTCGCAAGCTTCTCTTTACATCGAATCGTT 120
QY 123 GAGTTAACTCGGATTTGTTCTGCATCCTCGTTTCTGAAATCGGGGCATCTTAATTTG 182
Db 121 GAGTTAACTCGGATTTGTTCTGCATCCTCGTTTCTGAAATCGGGGCATCTTAATTTG 180
QY 183 TCTCGAATCTTCACCAATGCTTCGATCAAGCTGATGGTTAACAGTTGCTTAAG 242
Db 181 TCTCGAATCTTCACCAATGCTTCGATCAAGCTGATGGTTAACAGTTGCTTAAG 240
QY 243 ATCAGATCTTTGACCAAAATTTTGTGCTGATCTTTCAATTCGAAACGACACGCAA 302
Db 241 ATCAGATCTTTGACCAAAATTTTGTGCTGATCTTTCAATTCGAAACGACACGCAA 300
QY 303 ACAACCTCTGTATATGATTCGAGAAATTCGAAATTCGAAATTCGAAATTCGAA 362
Db 301 ACAACCTCTGTATATGATTCGAGAAATTCGAAATTCGAAATTCGAAATTCGAA 360
QY 363 AGACAGTTATTCGAAACACGAGTCAATGACGTGCGAAGAAACACCGAAGAAACATGAG 422
Db 361 AGACAGTTATTCGAAACACGAGTCAATGACGTGCGAAGAAACACCGAAGAAACATGAG 420
QY 423 CTTTCGAAGTTGCGCGAGAAAGCAAGCAAGATTACGCGGAGTTTCTCGCGCAGAA 482
Db 421 CTTTCGAAGTTGCGCGAGAAAGCAAGCAAGATTACGCGGAGTTTCTCGCGCAGAA 480
QY 483 AAGAGTAGATCCCGCGGAATTAATTCAGTTAGAGTCTCTGTTCTTAATGGTTGCT 542
Db 481 AAGAGTAGATCCCGCGGAATTAATTCAGTTAGAGTCTCTGTTCTTAATGGTTGCT 540
QY 543 AAGGAGATCTGTATTAATCTCCGTTAGTTTGGCGCGTACTCCGCTCCGTTAT 602
Db 541 AAGGAGATCTGTATTAATCTCCGTTAGTTTGGCGCGTACTCCGCTCCGTTAT 600
QY 603 GCTGCTATTGTTCCCATTTGAGATTAAGTTTACTCGTCTGTTCTGCTTCATGCTTC 662
Db 601 GCTGCTATTGTTCCCATTTGAGATTAAGTTTACTCGTCTGTTCTGCTTCATGCTTC 660
QY 663 TTTTCATTTCAATGCTAATACAGATGTAATCTTGCAATTTCAATGAAATTTAT 722
Db 661 TTTTCATTTCAATGCTAATACAGATGTAATCTTGCAATTTCAATGAAATTTAT 720
QY 723 ATATGATGAT 732

Db 721 ATATGATGAT 730

RESULT 4
LOCUS BT011724
DEFINITION Arabidopsis thaliana At3g59900 gene, complete cds.
ACCESSION BT011724
VERSION BT011724.1 GI:44917524
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae; eurosida II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shin,P., Carninci,P., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Narusaka,M., Sakurai,T., Satou,M., Seki,M., Shinozaki,K. and Becker,J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 393)
TITLE Cheuk,R., Chen,H., Kim,C.J., Shin,P., Carninci,P., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Narusaka,M., Sakurai,T., Satou,M., Seki,M., Shinozaki,K. and Becker,J.R.
JOURNAL Direct Submission
SUBMITTED (04-MAR-2004) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PONT (ORF) clones using the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Shin,P., Bowser,L., Chan,M.M., Chang,C.M., Dale,J.M., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Davis,R.W., Theologis,A., and Becker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Becker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES
SOURCE location/Qualifiers
1..393
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U5594"
/ecotype="Columbia"
1..393
/note="unknown protein"
/codon_start=1
/product="At3g59900"
/protein_id="AA549087.1"
/db_xref="GI:44917525"
/translation="MIREISNLQKDIINIDSYNNRVMVDGNNRKNKMSFRSSPEKS
KQELRSPSAQRMMIPANYPSSLSLFLVLTASLILPLVLPPLPPPLMLLVPI
GIVTLVLAIPSSSHNATDVICNFW"

ORIGIN
Query Match 53.7%; Score 393; DB 15; Length 393;
Best Local Similarity 100.0%; Pred. No. 1,4e-89;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

316 ATGATTCGAGAAATCTCAAACTTACAAAAGATATTATTAACATTCAGACAGTTATTCG 375

Db 1 ATGATTCGAGAAATCTCAAACTTACAAAAGATATTATTAACATTCAGACAGTTATTCG 60

Qy 376 AACCAACGAGTCATGACGTCGGAAAGAACACCGGAAAAACATGAGTTTCGAAGTTCCG 435

Db 61 AACCAACGAGTCATGACGTCGGAAAGAACACCGGAAAAACATGAGTTTCGAAGTTCCG 120

Qy 436 CCGGAGAAAAGCAAGCAAGATTACGGCGGAGTTTCGCGGCGAGAAAAGATGATGATC 495

Db 121 CCGGAGAAAAGCAAGCAAGATTACGGCGGAGTTTCGCGGCGAGAAAAGATGATGATC 180

Qy 496 CCGGCAATTAATTCAGTTAGAGTCTGTTCCTATTTGGTGTCTAACGCAATCTCTG 555

Db 181 CCGGCAATTAATTCAGTTAGAGTCTGTTCCTATTTGGTGTCTAACGCAATCTCTG 240

Qy 556 TTAATACCTCCGTAGTTTTCGCCGCCGTACCTCCGCTCCGTTATGCTCATTTGTT 615

Db 241 TTAATACCTCCGTAGTTTTCGCCGCCGTACCTCCGCTCCGTTATGCTCATTTGTT 300

Qy 616 CCCATTGGGAAATTAATGTTTACTGCTGTTCTTCGCTTCATGACCTTCCTCATTTAAT 675

Db 301 CCCATTGGGAAATTAATGTTTACTGCTGTTCTTCGCTTCATGACCTTCCTCATTTAAT 360

Qy 676 GCTAATACAGATGTACTGCAATTTTCATGTAA 708

Db 361 GCTAATACAGATGTACTGCAATTTTCATGTAA 393

RESULT 5
AY080817
LOCUS
DEFINITION Arabidopsis thaliana unknown protein (At2g44080) mRNA, complete cds.
ACCESSION AY080817
VERSION AY080817.1 GI:19423989
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae; eurosida II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Yamada,K., Liu,S.J., Sakano,H., Pham,P.K., Banb,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shin,P., Southwick,A., Shinozaki,K., Davis,R.W., Becker,J.R. and Theologis,A.
TITLE Arabidopsis Full Length cDNA Clones
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 904)
TITLE Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,B., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyer,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shin,P., Southwick,A., Shinozaki,K., Davis,R.W., Becker,J.R. and Theologis,A.
JOURNAL Direct Submission
SUBMITTED (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,B., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,

Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,
Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P.,
Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.

FEATURES

Source

Location/Qualifiers

1..904

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="2"

/clone="RAFL06-04-F13 (R11991)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector
(Lambda ZAP) as a XhoI/SacI insert."

1..904

/gene="At2g44080"

1..443

/gene="At2g44080"

444..851

/gene="At2g44080"

/codon_start=1

/evidence=experimental

/product="unknown protein"

/protein_id="AAL87296.1"

/db_xref="GI:19423930"

/translation="MIREFSLQNDIINIOEHYSLNNMVDGRHNRKNTSPFGSAPA
PIHGKELPRTLSQNSPRLISASYSMLVGLTASLILPLILPLPPPPPM
LILIPGIMVLMLVLAIFMPSNRKHSSTPM"

852..904

/gene="At2g44080"

ORIGIN

3' UTR

Query Match 24.9%; Score 182.6; DB 15; Length 904;
Best Local Similarity 70.0%; Pred. No. 9.2e-36;

Matches 296; Conservative 0; Mismatches 109; Indels 18; Gaps 3;

```

QY 270 CTGATCTTCAATCAACACAGACAGCAAAACAACCTCTGATGATTCAGAGAAT 329
    |||||
DB 398 CTGATTTTCAGATCAAACTCGAAAAGAGAAAAGCTCTTAATGATTCGTGAGTT 457
    |||||
QY 330 CTCAAATTTACAAAAGATTTAATCAATTCGAAGACATTCGAACACGAGTCAT 389
    |||||
DB 458 CTCAGATCTCAAAAACGACATCAATTCGAAGACATTCCTCAACACACAT 517
    |||||
QY 390 GGAAGTGGAGA---AACAAACGGAACAAACATGAGCTTCGAGTT-----CGCC 437
    |||||
DB 518 GGAAGTGAAGAGATCATTAACGGAACAAACAGATTTCTGAGTTCAGCTCCAGCTCC 577
    |||||
QY 438 GGAAGAAAGCAAGACAGATTACCGCGAGTTTCTCGCGCAGAA-----AAGATGAT 491
    |||||
DB 578 GATTATGGGGAAGAGATTGTTTCGACACTTCGTGCGAGAACAGTCCAGAGAGGCT 637
    |||||
QY 492 GATCCCGGCAATATTTCAGTTAGAGTCTCTGTTCTTAATGTTGTTGTTCAACGCGATC 551
    |||||
DB 638 AATATACGCAATCTCACTTCAATCAATCAATGTTGTTGTTGTTGTTGTTCAACGATC 697
    |||||
QY 552 TCTGTTAATCTTCGTTAGTTTGGCGCGTTCCTCCGCGCTCCGTTAATGCGCATTT 611
    |||||
DB 698 TCTCTGATCTTACCGTTGATTTCTTCACCAATTCCTCCTCCTCTTTTAAGTCGCTTTT 757
    |||||
QY 612 GGTTCATTTGAGATTATGTTTACTGTCGTTCTTGCTTCAGTCCCTTCCTTCATTC 671
    |||||
DB 758 GATTCATTTGAGATTATGTTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 817
    |||||
QY 672 TAA 674
    ||
DB 818 CAA 820

```

```

RESULT 6
AC004005
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
TITILE
REFERENCE
AUTHORS
JOURNAL
TITILE
COMMENT
FEATURES
source
1..109741
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="2"
/map="C1C10F02, C1C02B07"
/ecotype="Columbia"
/complement(1..4149)
/note="overlap with BAC clone F18019 (AC002333:1..4149)."
/rpt_family="AT_rich"
3302..4750
/gene="At2g43880"
/note="synonym: F6E13.1"
join(3302..3748,3846..4135,4219..4447,4532..4750)
/gene="At2g43880"
join(3302..3748,3846..4135,4219..4447,4532..4750)
/gene="At2g43880"
/codon_start=1
/product="putative polygalacturonase"
/protein_id="AAC23426.1"
/db_xref="GI:3212875"
/translation="MANNISFPCAIIFPSINIFLKSHPAMPFNVORGARDDGRA
DAKSTRTANSLACGRARARVYVPGTYIVKLVYMGPKNIIITKNDITLYAPANY
WDIGNSGQWTLFAKVRISYVGGTIDARAGVIRSGCRKSGHCHQAGARSISFVSCNVL
LSGLSFNSGNMHWIVHSSNRIENVRAPSGSPPTDIIHVOSSSGVITSGTQAT
GDCIALSQSGSNMIVIERVNCGPGHGISISLSDVYANBEQVNVTVLSVFTKQNV
RITMYRPSRGPFWNIVFRLLIMNNVENPYLIQNYCPNKGCPROSGVKGISGVTA
NKGSTTPIAMLDGSGSNHCTGLRLODKLTYMRSASYSYCRNMGRAAGVWVPPN
CM"
complement(5593..5613)
/rpt_family="AT_rich"
/rpt_family="AT_rich"
complement(6730..6753)
/rpt_family="AT_rich"
7192..8658
/gene="At2g43890"
/note="synonym: F6E13.2"
join(<7192..7632,7743..8032,8122..8350,8440..8658)
/gene="At2g43890"
join(7192..7632,7743..8032,8122..8350,8440..8658)

```


Db 68067 GAGCGTGAAGAGATCATTAACCGGAAAAACAGAGTTTCTGCGTTCAAGCTCCAGCTCC 68126

Qy 438 GGAGAAAAGCAAGCAAGATTACCGGAGTTTCTCGCGCAGAA-----AAGGATGAT 491

Db 68127 GATTATGGGAGAGCAAGATTGTTTCCGACATTCTCGTCAGAACAGTCCACAGAGGCT 68186

Qy 492 GATCCCGGCAATATTATTAGTTAGTCTGTTCTCTATTGTTGGTTGCTTAACGCGATC 551

Db 68187 AATATGACGCGAGTACTTCACTTGAATCAATGTTGTGCTTGTGTCTCAAGATC 68246

Qy 552 TCTGTTAACTCTCCGTTAGTTTCCGCGGTTACCTCCGCTCGTTTAACTGCTATT 611

Db 68247 TCTCTGATCTTACCGTTGATCTTCCACCATTCGCTCTCTCTTTTAACTGCTATT 68306

Qy 612 GGTTCGATGAGATTAAGTTTACTCGTGTCTTGCCTTAAGCTTCTCTGATTC 671

Db 68307 GATTCTATTGGAGTTAAGTTTCTTAACTGTTCTTAACTGCTTAACTGCTTAACTC 68366

Qy 672 TAA 674

Db 68367 CAA 68369

RESULT 7
BT000930 439 bp mRNA linear PLN 16-OCT-2002
LOCUS Arabidopsis thaliana clone Cl05228 unknown protein (At3g44080)
DEFINITION mRNA, complete cds.
BT000930
VERSION BT000930.1 GI:24030324
KEYWORDS F1.1 CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 439)
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, B.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shin, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
Submitted (16-OCT-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
Annotation based on January 2002 version of the Arabidopsis genome submitted to Genbank.
TITILE Arabidopsis Open Reading Frame (ORF) Clones
REFERENCE JOURNAL Unpublished
AUTHORS 2 (bases 1 to 439)
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, B.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shin, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (16-OCT-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
Annotation based on January 2002 version of the Arabidopsis genome submitted to Genbank.
COMMENT JOURNAL
FEATURES
source 1. .439
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="Cl05228"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1. .439
/gene="At2g44080"
1. .408
/gene="At2g44080"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA041330.1"
/db_xref="GI:24030325"

ORIGIN
Query Match 22.8%; Score 167; DB 15; Length 439;
Best Local Similarity 71.4%; Pred. No. 9.4e-32;
Matches 269; Conservative 0; Mismatches 90; Indels 18; Gaps 3;
3' UTR
/translacion="MIREFSLQNDIINIIOEHSYLNNDVRDHRKNTSPGSA
PIMKQELPRITLSSSPRLISASYSLSMVVLVGLTASLLILPLILPLPLPPPM
LILIPGIVULMVLAFMPSNSKHSSTFM"
409. .439
/gene="At2g44080"

Db 316 ATGATTCGAGAAATCTCAAACTTACAAAAGATATTAACATTCAGACAGTATTTCG 375

Qy 1 ATGATTCGAGATTCCTCCAGCTTACAAAAGATCAATTAACATTCAGACAGTATTTCG 60

Db 376 AACACCGAGTACTGACAGCTCGGAAG--AACACCGGAAAAACATGACTTTCGAAGT 432

Qy 61 CTCACCAACAACATGACGAGAGATCATACCGGAAAAACAGAGTTTCTGCTGT 120

Db 433 T-----CGCCGAGAAAACAGCAAGATTCAGCGGAGTTTCTCGCGCAGAA- 482

Qy 121 TCAGCTCCAGCTCCGATTATGGGAGACAGAAATGTTTCGACATTGTCTGCGAGAAC 180

Db 483 -----AAGATGATGATCCCGGAAATTATTCACTTGAAGTCTCTGTTCTATTGGTT 537

Qy 181 AGTCAAGAGAGGCTTAATATACGAGATTAATCAATCAATGTTGCTGTT 240

Db 538 GGTCTAACGAGCATCTCTGTTAACTTCCGTTAGTTTCCCGGCTTACTCCGCTCCG 597

Qy 241 GGTCTACAGACATCTCTGATCTTACCTTATCTTCCACATTCGCTCTCCCTCT 300

Db 598 TTTATGCTGCTATTGTTGTTCCATTTGGATTAAGTTTACTCGTCTGCTTTCGCTCATG 657

Qy 301 TTTATGCTGCTATTGTTGTTCCATTTGGATTAAGTTTACTCGTCTGCTTTCGCTCATG 360

Db 658 CCTTCTTCTGATTTCTAA 674

Qy 361 CTTCTTCTTATTTCCAA 377

RESULT 8
AX506572 336 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 1267 from Patent WO0216655.
DEFINITION AX506572
ACCESSION AX506572
VERSION AX506572.1 GI:23387809
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
1
Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing same, and methods of use
Patent: WO 0216655-A 1267 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG (CH)
TITILE Arabidopsis thaliana
REFERENCE JOURNAL
AUTHORS 1. .336
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 18.1%; Score 132.6; DB 6; Length 336;
Best Local Similarity 71.3%; Pred. No. 6e-23;
Matches 209; Conservative 0; Mismatches 69; Indels 15; Gaps 2;
Db 397 GGAAGAAACAACCGGAAAAACATGAGCTTTCGAGTT-----CGCCGAGAAAAAGC 447

Qy 13 GGAGATCATTAACCGGAAAAACAGAGTTTCTGCGTTCAAGCTCCAGCTCCGATTAAGGAG 72

and Genome Science Laboratory in Riken, Adachi, Y., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Heshikume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kageawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komio, H., Kouda, M., Koye, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishik, K., Nomura, K., Numasaki, R., Onno, M., Oosato, N., Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasakidi, D., Sato, K., Shibata, C., Shinagawa, A., Shirai, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Kenhta, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashitaki, Y.

FEATURES	Location/Qualifiers
source	1. .1302

ORIGIN

Query Match	10.3%;	Score 75.6;	DB 15;	Length 1302;
Best Local Similarity	66.7%;	Pred. No. 2.1e-08;		
Matches 108; Conservative	0;	Mismatches 54;	Indels 0;	Gaps 0;

Qy	498	GCGGATATATTCAGTTTGAAGCTCTGTCTCTATTTGATGATCTAAACGCAATCGT	557
Db	865	GCGCAGATCTTCTCGATGAATCGATGTCTCTGTGATGATGCGAAGCGTGCCTGT	924
Qy	558	AAATCTTCGATGTTTTCGCGCGATACACTCGCATTCGATATGTTTC	617
Db	925	GATCTTCGCTGTCTGTGCGCGGTGCCCCGCGCGATGATGATGCTGTGTC	984
Qy	618	CATTGGATATGATTTTACTGTCGTCTTGCTTAAATGCC	659
Db	985	GATGGCGATGCTGTGCTCTCTGTGCGCTGACGTTAAATGCC	1026

RESULT 11					
OSJN00249	OSJN00249	46422 bp	DNA linear	PLN 16-APR-2005	
LOCUS	Oryza sativa genomic DNA, chromosome 4,	BAC clone:	OSUNBA0058G03,		
DEFINITION	complete sequence.				
ACCESSION	AL731606				
VERSION	AL731606.2	GI:32482924			
KEYWORDS	HTG.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group)				

REFERENCE	AUTHORS
1	Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J., Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., Tang, Y., Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L. S., Yu, Z., Fan, D., Li, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J., Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R., Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, X., Jia, J., Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q., Zhang, X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S., Ni, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J., Li, J., Hong, G., Xue, Y. and Han, B.

Sequence and analysis of rice chromosome 4
 Nature 420 (6913), 316-320 (2002)

TITLE
 JOURNAL

TITLE	JOURNAL	PUBMED	REFERENCE
Sequence and analysis of rice chromosome 4	Nature 420 (6913), 316-320 (2002)	12447439	2
<p>Han, B., Feng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q., Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Weng, Q. J., Zhang, L., Lu, Y. Q., Yu, S. T., Liu, X. H., Lu, T. T., Zhang, Y. T., Lu, Y. X., Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K., Zhou, B., Chen, Z. H., Hao, P., Zhang, L., Wu, M., Zhang, R. Q., Guan, J. P., Fu, G., Wang, S. X., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F., Tu, Y. F.</p>			

REMARK	JOURNAL	TITLE
Direct Submission	Submitted (04-MAY-2002)	Han Bin, National Center for Gene Research
	Chinese Academy of Sciences, 500# Cao Bao Road, Shanghai 200233, CHINA. E-mail enquiries: bhan@ncgr.ac.cn	
	bhan@ncgr.ac.cn	
	Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: OSJNB00058G03.	
	On Jul 9, 2003 this sequence version replaced gi:20451715.	

Genes were identified by a combination of several methods: Gene prediction programs including *genseh* (<http://www.softberry.com/>), *genescan* (<http://CCR-081.mit.edu/GENSECAN.html>), *GeneMarkES* (<http://genemark.biology.gatech.edu/GeneMark/>), *GeneMark-ES* (<http://genome.wustl.edu/eddy/eddy/CRNscan-ES/>), searches of the complete sequence against NCBI non-redundant protein database (<http://ncbi.nlm.nih.gov/blast/d/>) and the EST database at NCGR.

FEATURES	Location/Qualifiers
source	1. .46422

```

gene      17958..18485
          /gene="OSJNBAA00586G03..3"
CDS       17958..18485
          /gene="OSJNBAA00586G03..3"
          /codon_start=1
          /protein_id="CAB02428..1"
          /db_xref="GI:32483262"
          /db_xref="GOA:Q7XRML"
          /db_xref="InterPro:IPR007087"
          /db_xref="UniProt/TREMBL:Q7XRML"

```

ORIGIN	Query Match	10.3%;	Score 75.6;	DB 15;	Length 46422;
	Best Local Similarity	66.74;	Pred. No. 1.7e-08;		
	Matches 108;	Conservative 0;	Mismatches 54;	Indels 0;	Gaps 0;
Qy	498	GGCGAATTAATTCAGTTAGAGTCTCTGTTCTTAATGTTGGTCTTACAGGCATCTCTGTT	557		
Db	32928	GGGCAAGTACTTCTCGTGGAAGTGGCTGCTCTGCTGGTGGTGGCGAGAGGGGCTGCTGT	32987		
Qy	558	AATACTTCGGTTAGTTTGGCCGCGGTACTCTCGCCTCCGTTTATGCTGCTATTGGTTCC	617		
Db	32988	GATCCCTCCGCTCTGCTGCTGCGCGCGCTGCCCCCGCGCGCTGCGAATGCTGCTGCTCC	33047		
Qy	618	CATTGGATTAATGATTTTACTGCTGCTGCTTTCCTTCATAGCC	659		
Db	33048	GGTGGAGTCTGGTGGTCTGCTGCTGCGGCGCTTCAGCC	33089		
RESULT 12					
AC157500					
LOCUS	AC157500	90825 bp	DNA	linear	PLN 18-FEB-2005
DEFINITION	Oryza sativa (japonica cultivar-group) BAC clone OSUNB0056004,				
ACCESSION	AC157500				
VERSION	AC157500.1	GI:59933327			
KEYWORDS	HMG.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group)				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				
AUTHORS	1 (bases 1 to 90825)				
	McCombie,W.R., Zuluvern,T., Nascimiento,L., Balija,V., Kramer,M., Spiegel,L., Miller,B., Muller,S., Katzenberger,F., Andrade,M.V., Palmer,L. and Bell,M.				
	Oryza sativa (japonica cultivar-group) BAC clone OSUNB0056004, from chromosome 4, complete sequence				
	Unpublished				
	2 (bases 1 to 90825)				
	McCombie,W.R.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (18-FEB-2005) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA				
AUTHORS	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. Clone OSUNB0056004 overlaps clone OSUNB0027P08 (AL731593) from base 54646 to base 90825. The overlap is from base 1 to base 36180 on OSUNB0027P08.				
JOURNAL	Location/Qualifiers				
	1. 90825				
COMMENT	/organism="Oryza sativa (japonica cultivar-group)"				
	/mol_type="genomic DNA"				
	/cultivar="Nipponbare"				
	/db_xref="taxon:39947"				
	/chromosome="4"				
	/clone="OSUNB0056004"				
	/clone_11b="EcoRI"				
	36680..36760				
	/note="The assembly is covered by high quality sequences derived from a transposed plasmid subclone."				
	37336..37516				
	/note="The assembly is covered by high quality sequences derived from a transposed plasmid subclone."				
	41013..41445				
FEATURES					
source					
	misc_feature				
	misc_feature				
	misc_feature				
	misc_feature				

Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (<http://www.softberry.com/>), genscan (<http://CCR-061.mt.edu/Genscan.html>), GenemarkHM (<http://genemark.biology.gatech.edu/Genemark/>), tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>), searches of the complete sequence against NCBI non-redundant protein database (<http://nbl.nlm.nih.gov/biast/db>) and the EST database at NCGR.

Location/Qualifiers

1. 118959
/organism="Oryza sativa (japonica cultivar-group)"

source

/mol_type="genomic DNA"

/cultivar="japonica"

/cultivar="Nipponbare"

/sub_species="japonica"

/db_xref="taxon:39947"

/chromosome="4"

/clone="OSUNBA0027P08"

/clone_id="CG1-OSUNBA"

21. 452

/gene="OSUNBA0027P08.1"

21. 452

/codon_start=1

/protein_id="CAD40976.1"

/db_xref="GI:21742738"

/translation="MTLALPTQOONFEFSALGSGMOPSLPFPDMSRLQENRGVDD
YLMREIRVRSHEILENEDMOHMLRLISMGSGSNHMGDSFPMSPAPARYBDD
RARPSSGAVVGMKIKAAHMGIFVRKKAERRAQLVELD"

4021. 4557

/gene="OSUNBA0027P08.2"

4021. 4557

/codon_start=1

/protein_id="CAD40975.1"

/db_xref="GI:21742737"

/translation="MYLISPRNGEDBDEQEIQLISDDEPPLKTLASCATASSSS
SGSDMEKGRGKACGGSTADPPPPSSSGSGGGGGSNIREAASGGGGMKRYFVS
ESLLILVCYRSLVILPLVPLPPLPPPSMLVPLVMLVLLALAPFTTSSSSAG
GGGGGNGATTHGAPYL"

complement (7119. 11761)

/gene="OSUNBA0027P08.3"

complement (join(7119. 7289,7595. 7655,7778. 7845,
8385. 8594,8728. 8768,8872. 8921,9121. 9209,9329. 9403,
9652. 9743,9992. 10043,10376. 10492,11128. 11357,
11479. 11533,11642. 11761))

/gene="OSUNBA0027P08.3"

/codon_start=1

/protein_id="CAD40974.1"

/db_xref="GI:21742736"

/db_xref="GOA:Q7X7KA"

/db_xref="InterPro:IPR004326"

/translation="MAATBAITTEDTPTWVAACSAIVLISAPFERSLHYLKALER
RRLVLEALKLKEMLLGIISLLVPEPQRIIASLIMGHMPCPSDKASIS
HGVAAVAASAVSGARLLGEGTAGSGSGKQVPLSLHAIIOHIFIVLALIT
ITIMRSFPKQFYSVTQDITAKLGVMEHFRGPKFNTYDMIKALEDKRYVGI
KMYIMFVMIPLFLNITGMSYFISLIPVLLILGTKEHLITQLAYVATKHAIV
EGDIAVSPDNLFWFHSRPLVLAIRLIFONAFEPYFTWATPGFNSCIDRLRY
RVSRIYICVVQVVCSTYSLPLVLAIVHMSGSSFKSAVSDVDADNLAKMDEARRRG
RAAGVGLGAAAGSSRRBGHIQNM"

14972. 16249

/gene="OSUNBA0027P08.4"

14972. 16249

/gene="OSUNBA0027P08.4"

/codon_start=1

/protein_id="CAD02442.2"

/db_xref="GI:38344536"

/db_xref="InterPro:IPR000717"

/db_xref="UniProt/TREMBL:Q7X5X9"

/translation="MSSMESYLPATTESIAKQKADASESISILYRVLDPPSSAS

ALRTKELATNTLNTYLTENRAREELNLTOLRPPFSLPKAKTAKIVGIDAVAKI
PGTSEQLISLCKEMVETAEKRTFPRQVAKALALLENQYATLTLLTGLGV
RLDDLTLDVLDLLESKIFSLRNPJKASITARTARTALNAYVPAQGLDLSG
ILHAERKDYKTAYSYFEPAPFAFSAEDPKAFSLCKMLCKTMVQADVAQITISK
AGIKYIPDVAADAVSKRSLSFTSLAKDYQSULESDYVHSHLSLDTLLE
QNCRLIEPSRYEIAHIAEMIELPDVHEKLSQMLIDKPKAGTLDQAGCULIFED
PKTEAFPLATLETISNVGKVDLSYKRSAKINA"

gene

CDS

/complement (21744. 23153)

/gene="OSUNBA0027P08.5"

/complement (21744. 23153)

/gene="OSUNBA0027P08.5"

/codon_start=1

/protein_id="CAD40973.2"

/db_xref="GI:38344537"

/db_xref="GOA:Q7X6H5"

/db_xref="InterPro:IPR000051"

/db_xref="UniProt/TREMBL:Q7X6H5"

/translation="MHTSTKLRTPTQOOLDAALINDGHRSPAPRPAPRSKML
LLVIATNLVSYLFGSGLSLRLPAGAAAPSHLMDSSALLRDLGTRALAAARAV
AALRAQCNASLLLSVLAGLGAHDGKPAADRGFGPPEPTGLRLATEPHRLPL
GSAKLGTDLEHGVGPAACRNQDELARVMAVDAGRCDDADALSLQILKGCERLP
RRRCRPSRPARVYEPAPLPGSLMSTPDTTWNVSPACKNYCTLVRAARAGGCGSY
EKCDCFDLAGKERRRMSDNGRPFSDGVLASAPGVGVLIGGAGGFLAAM
RERGVTVTTLQVGAFFSAFVASRGLVBLQSLAQLPLADGMDIYHMQIGVVP
GAVLELALPDVYVLRPGVFMLDHFACVCPRLNDTYABTLDRVGFRRLRWKSRXLD
LGAERREYLSALLERPLT"

28302. 32147

/gene="OSUNBA0027P08.6"

join(28302. 28530,28643. 28739,29125. 29322,29768. 30002,
30244. 30571,30659. 30929,31030. 31285,31369. 31825,
31957. 32147)

/gene="OSUNBA0027P08.6"

/codon_start=1

/protein_id="CAD40972.2"

/db_xref="GI:38344538"

/db_xref="GOA:Q7XUW2"

/db_xref="InterPro:IPR007778"

/db_xref="UniProt/TREMBL:Q7XUW2"

/translation="MAQBERGRLDQGNAAAGGAGWPPALARTVLYVMAYVLT
MAAIFPTFPAQAPBGMATKOSLFGITGSIFLAPSLPILVLAALVYVTAAP
SDIHVKNKRLKTLCFRLMTFPLVDGPPGVSAVEFGLVPLVYVVMYTAAP
VSLISKFGQISLTSSELILYIGLRGSLGFCMAFLPLVSRGSLVLRLLIDPFA
TRHWVLGHTMALFTLHGCVIYANSLBGNLGLAAMREISVANLPGVLSAAGL
LMWTSLAPRYKYFELPFTTHOYIIFVFLAAYVDDFISPSAGGIFLPMIDRPLF
MGRBATVDIISASCRGCVYELVFSRPAELRYALASFTYQVRELSFLMHPRSVGS
PMGRYHMSLILVDSWTKRGITTDQOSGRNSSTGTGILACVBPYHBSFY
HLMYENLIVAGIGISPLALISDIHRIEGRPCMPKRVLLVMKSNELSLGA
VDAQPTSSSVSDKLHIDQAFVQESQPLREGLVDDQKATGMFYKNGITVSGLVGT
GNFMAGMYFAASTLGFVLAVALAAYVRLVFMHMLGLVFLCMAAGVALLPGIL
VYLIMHLSERKAEEDRMDAALAAVPRADGGERAROTLNGDAAGGVSLAVKTR
YGRCPRPABAPFAFAKAGGAADVGLVCGPFGQASVARECRSHNLGRGRRRRA
GAVFHNHSPDI"

34750. 37049

/gene="OSUNBA0027P08.7"

join(34750. 34661,34996. 35028,35853. 36081,36570. 36687,
36771. 37049)

/gene="OSUNBA0027P08.7"

/codon_start=1

/protein_id="CAD40971.2"

/db_xref="GI:38344539"

/db_xref="GOA:Q7XUW3"

/db_xref="InterPro:IPR001965"

/db_xref="UniProt/TREMBL:Q7XUW3"

/translation="MDGYSVTVVHADASPBEVDPFCGRSGIYKALTIVYKTYK
QCDPERKENCILYGLPNGTAVVTLPADVEPELEPLAGINFADGQKQKMLLNAVH
SSGMLSLVAFYFGARPGFDKAEERLPMFTTSPLTYFEVVSQGVNTQOSKANSQK
SGSKPRRPNSSDKPOKOVAKYEENGRSGNGDDQATFICAGCEGAAYANSEFWIC
CDICETWFGKCYRITPAKAEHKKHYKCCGCKSKRTRE"

complement (37641. 42904)

/gene="OSUNBA0027P08.8"

complement (join(37641. 37734,37839. 38133,38241. 38316,
38526. 38714,38816. 39037,39176. 39274,39363. 39443,

gene

CDS

complement (37641. 42904)

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

```
39614. .39912,41472. .41811,41998. .41971,42052. .42273
42726. .42904))
/gene="OSJBA0027P08.8"
/codon_start=1
/protein_id="CAD40970.2"
```

Query Match	10.3%	Score 75.6;	DB 15;	Length 118959;
Best Local Similarity	66.7%;	Pred. No. 1.6e-08;		
Matches 108; Conservative	0;	Mismatches 54;	Indels 0;	Gaps 0;

QY 498 GGCAGATTATTCAGTTAGTAGAGTCTCGTTCCTCATTTGGTTGGTGTCAACGGCATCTCGTT 557

Db 4308 GGGCAGATCTCTTCGCTGAGTAGTGTCTCTCTGCTGTATGTGCTGACGGAGTGTGCTGT 4367

QY 558 AATATCTTCGTTAGTTTGGCCGCGGTAACTCCGCGCTCGTTTAGCTGCTATTTGTTCC 617

Db 4368 GATCTCTCCGCTGCTGTGCGCCGCGCTGCCCCGCGCGCTGAGTGTGATGCTGATGCTGTGTC 4427

QY 618 CATTGGATTATGTTTACTCGTGTCTTTCGCTTCAATGCC 659

Db 4428 GGTGGCGATGCTGTGCTGCTGTGCGGCTTGAGCTTATATGCC 4469

RESULT 15

LOCUS	DD099057	629 bp	mRNA	linear	PLN 18-JUL-2005
DEFINITION	Arachis stenosperma clone AS1RN2P12 microsatellite sequence.				
ACCESSION	DD099057				
VERSION	DD099057.1	GI:70779734			

SOURCE ORGANISM	
Arachis stenosperma	
Arachis stenosperma	

Bakarycea; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosid I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.

REFERENCE
AUTHORS
1 (bases 1 to 629)
Moretzsohn, M.C., Leol, L., Proite, K., Guimaraes, P.M.,

TITLE A microsatellite based, gene-rich linkage map for the AA genome of

JOURNAL Theor. Appl. Genet. (2005) In press
REFERENCE 2 (bases 1 to 629)

AUTHORS Bertoli, D.J.

TITLE Direct Submission
JOURNAL Submitted (17-JUN-2005) Laboratorio de Interacoes Planta-Pragas II

FEATURES

Location/Qualifiers

```
source
1. 629
/organism="Archaeis stenosperma"
/mol_type="mRNA"
/db_xref="taxon:217475"
/clone="ASINR212"
repeat_region
1. 629
/note="microsatellite"
/ptc_type=tandem
```

ORIGIN

Query Match	9.6%	Score 70.2;	DB 15;	Length 629;
Best Local Similarity	63.2%;	Pred. No. 5.4e-07;		
Matches 108; Conservative	0;	Mismatches 63;	Indels 0;	Gaps 0;

	QY	Db	QY	Db	QY
	498	209	558	269	618
	GCGCAATATTTACAGTTTAAAGACTCTCTGTCCATTTGGTGGCTAACGCGCATCTCGTT	GATGGCTTGTATACAGTTTAAAGGATACATCTGATCCCTGGTGGTTCTTGCGACACGTGCATNGGT	AAATCTTCGGTATGTTTGGCGCGCGTACTCTCGCGCTCGGTTAATGCTGTATGGATTC	TGTTCTTTCGCGATGACTGCCGCGCGTACACCGCGCGCGCTGCTCTCTCTTCTTTTCC	CATTGGGATTAATGTTTAACTGTCGTGTTCTTGCGCTTCATGACCTTCTTTCTCA
	557	268	617	328	668

THIS PAGE BLANK (U8PT0)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:46:01 ; Search time 510 Seconds
(without alignments)
9565.794 Million cell updates/sec

Title: US-10-715-129-1

Sequence: 1 ttgtcttcctcattccctca.....gaaattatcatgatgat 732

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21:*

1: geneeqn1980s:*
2: geneeqn1980s:*
3: geneeqn2000s:*
4: geneeqn2001as:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004as:*
13: geneeqn2004bs:*
14: geneeqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	732	100.0	732	14	ADZ67285
2	182.6	24.9	1005	3	AAC53786 Arabidops
3	132.6	18.1	336	6	ABZ13462 Arabidops
4	132.6	18.1	336	8	ADA67786 Arabidops
5	109.8	15.0	650	3	AAC51918 Arabidops
6	70.4	9.6	298	11	ACL131088 Rice abio
7	67.6	9.2	990	13	ADK64606 Plant ful
8	67.6	9.2	1157	13	ADK10467 Plant ful
9	67	9.2	1137	14	ABE67601 Rice geno
10	60.6	8.3	665	13	ADK49328 Plant ful
11	59.6	8.1	1055	13	ADT18995 Plant cdn
12	56.6	7.7	978	11	ACL29210 Rice abio
13	56	7.7	2000	8	ABZ15156 Arabidops
14	56	7.7	2000	6	ADA69206 Arabidops
15	45.8	6.3	366	13	ADR64525 Cotton cd
16	44.4	6.1	936	13	ADK37027 Plant ful
17	43.8	6.0	939	13	ADT17295 Plant cdn
18	43.8	6.0	2000	11	ACL37108 Rice stre
19	41.6	5.7	2000	8	ADA71938 Rice gene

C 20	41.2	5.6	2000	8	ADA71938	ADA71938 Rice gene
C 21	41.2	5.6	6012	4	AAS46298	AAS46298 Tumour su
C 22	41.2	5.6	6012	6	ABK31187	ABK31187 Signal tr
C 23	41.2	5.6	6012	6	ABL70148	ABL70148 Chemical
C 24	40.6	5.5	127145	13	ADQ80254	ADQ80254 Hermansky
C 25	40.2	5.5	7329	4	AAS46674	AAS46674 Tumour su
C 26	40	5.5	5559	13	ADR89733	ADR89733 Oligonuc
C 27	39.6	5.4	612	13	ADR72392	ADR72392 DNA encod
C 28	39.6	5.4	1717	6	ABQ75360	ABQ75360 Human lun
C 29	39.6	5.4	4412	13	ADR72391	ADR72391 DNA encod
C 30	39.4	5.4	450	13	ACN49835	ACN49835 Cotton pr
C 31	39.4	5.4	631	12	ADJ10666	ADJ10666 Recombina
C 32	39.2	5.4	19112	13	ADM48345	ADM48345 Marburg v
C 33	38.6	5.3	5429	4	AAS46590	AAS46590 Tumour su
C 34	38.6	5.3	5429	6	ABL33845	ABL33845 Human imm
C 35	38.6	5.3	9884	6	ABK33925	ABK33925 Human dna
C 36	38.6	5.3	9884	8	ADA20340	ADA20340 Prostate
C 37	38.6	5.3	9884	8	ADA64147	ADA64147 Human ren
C 38	38.6	5.3	24259	4	AAS46692	AAS46692 Tumour su
C 39	38.4	5.2	2000	11	ACL35363	ACL35363 Rice stre
C 40	38.4	5.2	2651	12	ADL13321	ADL13321 Human ste
C 41	38.2	5.2	26997	4	AAS46748	AAS46748 Tumour su
C 42	37.8	5.2	3095	14	ADM10508	ADM10508 Colon pro
C 43	37.6	5.1	579	13	ACN51616	ACN51616 Cotton an
C 44	37.6	5.1	7133	4	AAS46388	AAS46388 Tumour su
C 45	37.4	5.1	599	13	ACN52180	ACN52180 Cotton an

ALIGNMENTS

RESULT 1	ADZ67285	standard; cdna; 732 BP.
ID	ADZ67285	
AC	ADZ67285;	
DT	28-JUL-2005	(first entry)
XX	Arabidopsis thaliana auxin-inducible gene ARGOS.	
XX	cloning; plant breeding; auxin; development; transformation; gene; ss.	
XX	Arabidopsis thaliana.	
XX	Key	Location/Qualifiers
XX	CDS	388..708
XX		/*tag= a
XX		/product= "ARGOS protein"
XX	US2005108793-A1.	
XX	19-MAY-2005.	
XX	18-NOV-2003; 2003US-00715129.	
XX	18-NOV-2003; 2003US-00715129.	
XX	(TEMA-) TEMASEK LIFE SCI LAB.	
XX	Hu Y, Xie Q, Chua N;	
XX	WPI; 2005-354827/36.	
XX	P-PSDB; ADZ67286.	
XX	GENBANK; AY305869.	
XX	New auxin-inducible gene ARGOS from Arabidopsis, useful for regulating	
XX	organ development including size control in transformed plants and plant	
XX	cells.	
XX	Claim 1; SEQ ID NO 1; 22pp; English.	
XX	The invention relates to an isolated nucleic acid, i.e. auxin-inducible	

CC gene, ARGOS, from Arabidopsis, comprising fully defined 732 base pair
 CC sequence given in the specification, that encodes the amino acid sequence
 CC comprising fully defined 106 amino acids also given in the specification,
 CC or comprising their antisense nucleotide sequences. The nucleic acid is
 CC useful for regulating organ development including size control in
 CC transformed plants and plant cells. This sequence corresponds to the cDNA
 CC for the novel auxin-inducible gene ARGOS.

XX SQ Sequence 732 BP; 199 A; 170 C; 125 G; 238 T; 0 U; 0 Other;

Query Match 100.0%; Score 732; DB 14; Length 732;

Best Local Similarity 100.0%; Pred. No. 1e-201;

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TTGCTCTCATTTCCCTACTAGTACTGTTTACACAGTTTCTGATCAACCAAC 60
Db 1 TTGCTCTCATTTCCCTACTAGTACTGTTTACACAGTTTCTGATCAACCAAC 60
QY 61 CAATACACAAGCTTCTCAAACTCTCACCCTCAAGCTTCTCTTACATCTGAATCG 120
Db 61 CAATACACAAGCTTCTCAAACTCTCACCCTCAAGCTTCTCTTACATCTGAATCG 120
QY 121 TTGAGTTAACTCGGATTTGTTTGCATCTCTGTTTCTGAATCTGCGCCATCTTAATT 180
Db 121 TTGAGTTAACTCGGATTTGTTTGCATCTCTGTTTCTGAATCTGCGCCATCTTAATT 180
QY 181 TGTCTGCAATTTCTCACCAATGCTTGATGATCAAGCTGATTTGTTACAGTTGCCCTAA 240
Db 181 TGTCTGCAATTTCTCACCAATGCTTGATGATCAAGCTGATTTGTTACAGTTGCCCTAA 240
QY 241 AGATCAGATCTTTGAGCAAAATTTTGTCACTGATCTTCAATCCAAACAGACAGCA 300
Db 241 AGATCAGATCTTTGAGCAAAATTTTGTCACTGATCTTCAATCCAAACAGACAGCA 300
QY 301 AAACCACTCTGTAGATGATTCAGAAATCTCAACTTACAAAAGATTTAATTAACATT 360
Db 301 AAACCACTCTGTAGATGATTCAGAAATCTCAACTTACAAAAGATTTAATTAACATT 360
QY 361 CAAGACAGTATTGCAACAACCGAGTATGAGCGTGGAGAAACAACCGGAAACATG 420
Db 361 CAAGACAGTATTGCAACAACCGAGTATGAGCGTGGAGAAACAACCGGAAACATG 420
QY 421 AGCTTTCGAAGTTCCGCGAGAAAAGCAAGAGTTTCTCGGCGCAG 480
Db 421 AGCTTTCGAAGTTCCGCGAGAAAAGCAAGAGTTTCTCGGCGCAG 480
QY 481 AAAAGAGTATGATCCCGGCAATTTTCACTTAAAGTCTGTTCTAATTTGTTGTT 540
Db 481 AAAAGAGTATGATCCCGGCAATTTTCACTTAAAGTCTGTTCTAATTTGTTGTT 540
QY 541 CTAACGGCATCTCTGTTAATCTCCGTTAGTTTGCAGCGGTTACCTCGGCGCTCGTT 600
Db 541 CTAACGGCATCTCTGTTAATCTCCGTTAGTTTGCAGCGGTTACCTCGGCGCTCGTT 600
QY 601 ATGCTGCTATGTTGTTCCATTTAGGTTTATGCTTTCATCGGTTCTTGCTTCATGCT 660
Db 601 ATGCTGCTATGTTGTTCCATTTAGGTTTATGCTTTCATCGGTTCTTGCTTCATGCT 660
QY 661 TCTTTCATCTTAATGCTAATACAGATGTAATCTGCAATTTCAATGTAATCTGAATTTA 720
Db 661 TCTTTCATCTTAATGCTAATACAGATGTAATCTGCAATTTCAATGTAATCTGAATTTA 720
QY 721 TTATATGATGAT 732
Db 721 TTATATGATGAT 732

```

RESULT 2
 AAC53386 ID AAC53386 standard; DNA; 1005 BP.

AC AAC53386;
 XX

```

DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 74346.
DE Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
KM promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 17-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.

```

```
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 24.9%; Score 182.6; DB 3; Length 1005;
Best Local Similarity 70.0%; Pred. No. 3e+42;
Matches 296; Conservative 0; Mismatches 109; Indels 18; Gaps 3;

QY 270 CTGATCTTCTAATCCAAACGACAGACAGAAACAACTCTGTGATGATTCAGAAAT 329
      |||||
Db 499 CTGATTTTCCAGATCAAACTTCGAAAAGAGAAAAGCCTTTTAAATGATTCGTGATT 558
      |||||
QY 330 CTCAACTTACAAAATAATTTAATCAATTCAGACAGATTCGAAACCAAGCAT 389
      |||||
Db 559 CTCAGTCTACAAAACGACATCATTAACATTTCAAGAAATTTCTCAACAAACAT 618
      |||||
QY 390 GGAAGTCGAGAGA---AACAACGGAAGAAACATGAGCTTCGAAGTT-----CGCC 437
      |||||
Db 619 GGAAGTGAAGAGATCATATACGGAAGAAACAGAGTTTGTGTGTCAGCTCCAGCTCC 678
      |||||
QY 438 GGAAGAAAGCAGCAGAGATTACGCGCGAGTTCTCGCGCAGAA-----AAGATGAT 491
      |||||
Db 679 GATTATGCGGAAGCAGAAATGTTTCCGACATTCCTCGCAGAACGTCACAGAGGCT 738
      |||||
QY 492 GATCCGCGCAATATTTAGTTAGTCTCTGTTCTCTATTGTTGTTGTTCTAACGGCATC 551
      |||||
```

```

Db      739 AATATCAGCAGTACTTCACTTGTAGATCAATGTTGCTGTTGCTCAAGCATC 798
Qy      552 TCTGTATATCTTCCGTTAGTTTGGCCGCTTACCTCGGCTCCGTTATNGCTGCTATT 611
Db      799 TCTCTTATCTTACCTGATTTCTTCCACCATTTGCTCTCTCTTTATGCTGCTTTT 858
Qy      612 GGTTCCTATGGGATTAATGTTTACTGCTGCTTCTTCCCTTCAAGCTTCTTCAATTC 671
Db      859 GATTCCTATGGGATTAATGTTTGTGTTATGTTCTTCTTCAATGCTTCTTCAATTC 918
Qy      672 TAA 674
Db      919 CAA 921

RESULT 3
AB213462
ID      AB213462 standard; DNA, 336 BP.
XX
AC      AB213462;
XX
XX      21-JAN-2003 (first entry)
XX
DE      Arabidopsis thaliana stress regulated gene SEQ ID NO 1267.
XX
KW      Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
XX      Arabidopsis thaliana.
XX
XX      MO200216655-A2.
XX
PD      28-FEB-2002.
XX
PF      24-AUG-2001; 2001MO-US026685.
XX
PR      24-AUG-2000; 2000US-0227866P.
PR      26-JAN-2001; 2001US-0264647P.
PR      22-JUN-2001; 2001US-0300111P.
XX
XX      (SCRI ) SCRIPPS RES INST.
PA      (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI      Harper JF, Krepe J, Wang X, Zhu T;
XX
DR      WPI; 2002-304127/34.
XX
PT      Identifying a stress condition to which a plant cell has been exposed and
PT      producing plants with increased tolerance to these abiotic stresses.
XX
PS      Claim 144; SEQ ID NO 1267; 577pp + Sequence Listing; English.
XX
XX      The invention relates to identifying a stress condition to which a plant
XX      cell has been exposed, comprising: (a) contacting nucleic acid
XX      representative of expressed polynucleotides in the plant cell with an
XX      array or probes representative of the plant cell genome; and (b)
XX      detecting a profile of expressed polynucleotides in the plant cell
XX      characteristic of a stress response. The method is useful in the
XX      production of transgenic plants, cells and seeds and in producing plants
XX      with increased tolerance to abiotic stress. The present sequence is that
XX      of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
XX      in methods of the invention. Note: The sequence data for this patent is
XX      not represented in the printed specification but is based on sequence
XX      information supplied to Derwent by the European Patent Office
XX
SQ      Sequence 336 BP; 72 A; 74 C; 68 G; 122 T; 0 U; 0 Other;
XX
Query Match      18.1%; Score 132.6; DB 6; Length 336;
Best Local Similarity 71.3%; Pred. No. 6,1e-28;
Matches 209; Conservative 0; Mismatches 69; Indels 15; Gaps 2;
Qy      397 GGAAGAAACACCGGAAAAACATGAGCTTTGGAAGTT-----CGCCGAGAAAAAGC 447
Db      13 GGAATCATTAACCGGAAAAACACGAGTTTCTGGTTCACTCCAGCTCCGATTAATGGGG 72

```

```

Qy      448 AACGAAGTTACGGCGAGTTTCTCGCGCAGAA-----AAGGATGATGATCCGGCG 501
Db      73 AACGAAGATTTGTTTGGACATTTGTGTGCGAAGAACGACCAATATACAGCG 132
Qy      502 AATTAATTCAGTTTAAGTCTGTCTCTTAATGTTGGTTAAAGGACATCTGTTAATA 561
Db      133 AGTTACTTCAGTTTAAGATCAATGATGTTGTTGTTGTTCAACGATCTCTTGATC 192
Qy      562 CTTCGGTTAGTTTGGCGCGCTTACCTCGCCCTCGTTATGCTGATTTGTTCCATT 621
Db      193 TTACCGTTGATTTCTTCACCATTTGCTCTCTCTTTAATGCTTTGATTTCCATT 252
Qy      622 GGAATTAAGTTTACTCGCTGCTTCCCTTCATGCTCTTCTCATTTGAA 674
Db      253 GGGAATTAAGTTTGTGCTTAATGTTCTTGTCTTCAATGCTTCTTCAATTCGA 305

RESULT 4
ADA67786
ID      ADA67786 standard; DNA, 336 BP.
XX
AC      ADA67786;
XX
XX      20-NOV-2003 (first entry)
XX
DE      Arabidopsis thaliana gene, SEQ ID 23.
XX
KW      Plant; bacterial infection; fungal infection; viral infection; ds.
XX
XX      Arabidopsis thaliana.
XX
XX      MO2003000898-A1.
XX
PD      03-JAN-2003.
XX
PF      22-JUN-2001; 2001MO-IB001105.
XX
PR      22-JUN-2001; 2001MO-IB001105.
XX
XX      (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI      Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX
PI      Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
XX
DR      WPI; 2003-175290/17.
XX
PT      Identifying at least one gene involved in plant resistance or response to
PT      pathogenic infection for conferring resistance or tolerance to a plant to
PT      bacterial, fungal or viral infection by determining or detecting plant
PT      gene expression.
XX
PS      Claim 6; SEQ ID NO 23; 899pp; English.
XX
XX      The present invention relates to a method (M1) for identifying genes
XX      involved in plant resistance or response to pathogenic infection. M1
XX      comprises identifying a gene whose expression is significantly altered in
XX      the incompatible interaction of plant gene expression relative to
XX      expression of the gene in an uninfected plant, in a mutant plant that
XX      does not express a gene associated with response to pathogenic infection,
XX      or in a corresponding incompatible or compatible interaction. (M1) is
XX      useful for conferring resistance to resistance or tolerance to a plant to
XX      bacterial, fungal or viral infection. The present sequence was used to
XX      illustrate the invention.
XX
SQ      Sequence 336 BP; 72 A; 74 C; 68 G; 122 T; 0 U; 0 Other;
XX
Query Match      18.1%; Score 132.6; DB 8; Length 336;
Best Local Similarity 71.3%; Pred. No. 6,1e-28;
Matches 209; Conservative 0; Mismatches 69; Indels 15; Gaps 2;
Qy      397 GGAAGAAACACCGGAAAAACATGAGCTTTGGAAGTT-----CGCCGAGAAAAAGC 447
Db      13 GGAATCATTAACCGGAAAAACACGAGTTTCTGGTTCACTCCAGCTCCGATTAATGGGG 72

```


PR	25-MAY-1999;	99US-0136021P
PR	27-MAY-1999;	99US-0136392P
PR	28-MAY-1999;	99US-0136789P
PR	01-JUN-1999;	99US-0137222P
PR	03-JUN-1999;	99US-0137558P
PR	04-JUN-1999;	99US-0137502P
PR	07-JUN-1999;	99US-0137724P
PR	08-JUN-1999;	99US-0138094P
PR	10-JUN-1999;	99US-0138540P
PR	10-JUN-1999;	99US-0138847P
PR	14-JUN-1999;	99US-0139119P
PR	16-JUN-1999;	99US-0139455P
PR	16-JUN-1999;	99US-0139453P
PR	17-JUN-1999;	99US-0139432P
PR	18-JUN-1999;	99US-0139454P
PR	18-JUN-1999;	99US-0139459P
PR	18-JUN-1999;	99US-0139641P
PR	18-JUN-1999;	99US-0139642P
PR	18-JUN-1999;	99US-0139463P
PR	18-JUN-1999;	99US-0139750P
PR	18-JUN-1999;	99US-0139763P
PR	21-JUN-1999;	99US-0139817P
PR	22-JUN-1999;	99US-0139899P
PR	23-JUN-1999;	99US-0140353P
PR	23-JUN-1999;	99US-0140354P
PR	24-JUN-1999;	99US-0140659P
PR	28-JUN-1999;	99US-0140823P
PR	29-JUN-1999;	99US-0140991P
PR	30-JUN-1999;	99US-0141287P
PR	01-JUL-1999;	99US-0141842P
PR	01-JUL-1999;	99US-0142154P
PR	02-JUL-1999;	99US-0142055P
PR	06-JUL-1999;	99US-0142330P
PR	08-JUL-1999;	99US-0142803P
PR	09-JUL-1999;	99US-0142920P
PR	12-JUL-1999;	99US-0142977P
PR	13-JUL-1999;	99US-0143542P
PR	14-JUL-1999;	99US-0143624P
PR	15-JUL-1999;	99US-0144005P
PR	16-JUL-1999;	99US-0144408P
PR	16-JUL-1999;	99US-0144086P
PR	19-JUL-1999;	99US-0144335P
PR	19-JUL-1999;	99US-0144331P
PR	19-JUL-1999;	99US-0144332P
PR	19-JUL-1999;	99US-0144333P
PR	19-JUL-1999;	99US-0144334P
PR	19-JUL-1999;	99US-0144335P
PR	20-JUL-1999;	99US-0144332P
PR	20-JUL-1999;	99US-0144632P
PR	22-JUL-1999;	99US-0144884P
PR	21-JUL-1999;	99US-0144814P
PR	21-JUL-1999;	99US-0145086P
PR	21-JUL-1999;	99US-0145088P
PR	22-JUL-1999;	99US-0145085P
PR	22-JUL-1999;	99US-0145089P
PR	22-JUL-1999;	99US-0145152P
PR	23-JUL-1999;	99US-0145154P
PR	23-JUL-1999;	99US-0145218P
PR	26-JUL-1999;	99US-0145224P
PR	26-JUL-1999;	99US-0145276P
PR	27-JUL-1999;	99US-0145913P
PR	27-JUL-1999;	99US-0145918P
PR	27-JUL-1999;	99US-0145919P
PR	28-JUL-1999;	99US-0145961P
PR	02-AUG-1999;	99US-0146386P
PR	02-AUG-1999;	99US-0146389P

PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147319P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 06-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 12-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158359P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.
 Query Match 15.0%; Score 109.8; DB 3; Length 650;
 Best Local Similarity 74.6%; Pred. No. 3.3e-21;
 Matches 138; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 QY 490 ATGATCCCGGCAATATTTCAGTTAGAGTCCTGCTCATTTGGTTGGCTCAAGCA 549
 DB 402 ATTATCAGGTTGATTCCTGTTTTTACACTCAATGCTTGCTTGCTTCACAGCA 461
 QY 550 TCTCTGTAACTTCCTGCTAGTTTGGCCGCGTCACTCCGCTCGTTATGCTGTA 609
 DB 462 TCTCTGTGATCTTACCGTATGTTTCACCAATGCTCTCTCCTTTATGCTGCTT 521
 QY 610 TTGGTTCCCATTTGGGATTTATGCTTTTACTGCTGCTTCTGCTTCAATGCTTCTCAT 669
 DB 522 TTGATTCCTATTGGGATTAATGTTTGTATGCTTATGCTTCAATGCTTCTTAAAT 581
 QY 670 TCTAA 674
 DB 582 TCCAA 586
 RESULT 6
 ACL1088/c
 ID ACL1088 standard; cDNA; 298 BP.
 AC ACL1088;
 AC 02-JUN-2005 (first entry)
 DT 02-JUN-2005 (first entry)
 XX
 DE Rice abiotic stress response related polynucleotide SEQ ID NO:9651.
 XX
 KM ss; abiotic stress tolerance; transgenic plant; plant; cereal;
 KM agriculture.
 XX
 OS Oryza sativa.
 XX
 PN MO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002MO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 24-AUG-2001; 2001US-0314662P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Krepes J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
 PI Moughamer T, Provart N, Rieke D, Zhu T;
 XX
 DR WPI; 2003-248011/24.
 XX
 PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 PT
 XX
 PS Disclosure; SEQ ID NO 9651; 89pp; English.
 XX
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulation.
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid

CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention

XX Sequence 298 BP; 77 A; 71 C; 87 G; 62 T; 0 U; 1 Other;

Query Match 9.6%; Score 70.4; DB 11; Length 298;
Best Local Similarity 63.7%; Pred. No. 6.5e-10;
Matches 107; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 495 CCGGCGCAATATTTAGTTAGTCTGTTCCATTAGTGGTGGTAAAGCATCTCT 554
Dd 222 CCTACAAATTAATCTGACGACGAGTGTGTGTATATATCTGTAACCGTCTCT 163
Qy 555 GTTAATATCTTCGTTAGTTTCCCGCGGTACTCCGCTCCGTTAATGCTGTAATGTT 614
Dd 162 GCTGATCTCTCTCTCATCTGCGCGCACGCGCCGACCTTCGATGCTTCTGCT 103
Qy 615 TCCCATTTGGGATATAGTTTACTGCTGCTTCTTCCCTTATAGCTTC 662
Dd 102 TCCATTAAGGCTCTCTGTGTGCTGATGATCTTATGCTTCAAGCTTC 55

RESULT 7
ADK64606
ID ADK64606 standard; cDNA; 990 BP.

XX ADK64606;
Dd 21-APR-2005 (first entry)

XX plant full length insert polynucleotide seqid 35449.
XX plant protectant; plant growth regulator; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/) LIU J.

XX (ZHOU/) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABA/) TABASKA J E.

XX (CAO/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to

XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or

XX pests, for conferring increased resistance to plant disease, or for

XX improving yield.

XX Claim 1; SEQ ID NO 35449; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a

CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 990 BP; 185 A; 296 C; 271 G; 238 T; 0 U; 0 Other;

Query Match 9.2%; Score 67.6; DB 13; Length 990;
Best Local Similarity 61.2%; Pred. No. 7e-09;
Matches 109; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 482 AAGAGATGATGATCCCGGCAATTAATTTAGTGTCTGTTCCATTAGTGTCTC 541
Dd 682 ACAAGAGATGATGATCCCGGCAATTAATTTAGTGTCTGTTCCATTAGTGTCTC 741

Qy 542 TAAAGCATCTCTGTTAATATCTCCGTTAGTTTGGCGCGGTACTCCGCTCCGTTA 601
Dd 742 TCAACGCTCTGCTGCTGATATCTCCGCTGCTGCTCTCCCTGCGCGCGCGCTGCG 801

Qy 602 TGTCTGTTAGTTGTTCCATTGAGTTAATGTTTACTGTTGCTTGTGCTTATGCTC 659
Dd 802 TGTCTGTTAGTTGTTCCATTGAGTTAATGTTTACTGTTGCTTGTGCTTATGCTC 859

RESULT 8
ADK10467
ID ADK10467 standard; cDNA; 1157 BP.

XX ADK10467;

XX 21-APR-2005 (first entry)

XX plant full length insert polynucleotide seqid 5042.

XX plant protectant; plant growth regulator; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/) LIU J.

XX (ZHOU/) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABA/) TABASKA J E.

XX (CAO/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.

XX Claim 1; SEQ ID NO 5042; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp://seqdata.uspto.gov/sequence.html;docID:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.

XX Sequence 1157 BP; 225 A; 348 C; 299 G; 285 T; 0 U; 0 Other;

XX Query Match 9.2%; Score 67.6; DB 13; Length 1157;
XX Best Local Similarity 61.2%; Pred. No. 7.5e-09;
XX Matches 109; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 482 AAAGATGATGATCCCGGCAATATTTCAGTTAGATCTGTTCTATTGGTTGTC 541
DB 790 ACAAGAAAGTCATCGCGTCCACCTACTTCAGCATCGGGGGTCTGTGCTGCGCC 849
QY 542 TAAACGATCTCTGTTAATCTTCGTTAGTTTCCGCGGTTACTCCGCTCCGTTA 601
DB 850 TCACCGTCTGCTGCTGATCTGCGCTGCTGCTCCCGCCGCGCGCTGCG 909
QY 602 TCGCTGATTTGTTCCCATGCGATTATGTTTACTGCTGCTGCTGCTGCTGCTGCTG 659
DB 910 TCGTGTGTGCGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 967

RESULT 9

XX AEB67601
XX ID AEB67601 standard; DNA; 1137 BP.

XX AEB67601;

XX 22-SEP-2005 (first entry)

XX Rice genome derived DNA sequence, SEQ ID 2746.

XX transcription; gene regulation; transgenic plant; RNA interference;
XX transformation; antibody; ds.

XX Oryza sp.

XX JP2005185101-A.

XX 14-JUL-2005.

XX 11-DEC-2002; 2002JP-00383870.

XX 30-MAY-2002; 2002JP-00203269.

XX

PA (DOKU-) DOKURITSU GYOSEI HOJIN NOGYO SEIBUTSU SH.
PA (SEIB-) SEIBUTSUKU TOKUETI SANGYO GIJUTSU.
PA (DOKU-) DOKURITSU GYOSEI HOJIN RIKAGAKU KENKUSH.
PA (KOKU-) ZH KOKUSAI KAGAKU SHINKO ZAIDAN.

XX Kikuchi H, Hayashizaki Y, Otomo Y, Matsubara K, Murakami K,
XX Kishimoto N, Sato K, Nagata T, Kawakami N, Yazaki J, Ishikawa M;
XX Doi K, Kawai J;
XX WPI; 2005-566181/58.

XX Novel DNA encoding transcription factor, derived from rice plant, useful
XX for obtaining transcriptional-regulatory regions in plant and for
XX producing modified plant.

XX Claim 1; SEQ ID NO 2746; 2928bp; Japanese.

XX The invention relates to a novel DNA sequence encoding a transcription
XX factor derived from a plant. The invention further comprises antisense
XX RNA sequences, ribozyme activity RNA, RNAi sequences, a vector,
XX transformed plant cells, antibodies and proteins, all related to the
XX novel plant DNA sequences of the invention. The novel DNA is preferably
XX derived from a rice genome database. The invention further provides a
XX method for determining the transcription regulatory regions of the rice
XX genome. The novel DNA is useful for controlling the expression of a gene
XX in a plant and for producing a modified plant with desired and different
XX characteristics. The plant DNA and method enables the acquisition of many
XX transcriptional-regulatory regions. This polynucleotide represents a DNA
XX sequence taken from a rice genome clone library for use in the invention.
XX Note: This sequence is not shown in the specification. It has been
XX retrieved from a sequence listing in electronic format from the Japanese
XX Patent Office. The invention claims DNA sequences of SEQ ID Nos 1 to
XX 28469 and encoded protein sequences of SEQ ID Nos 28470 to 56791,
XX however, the sequence listing only provided the DNA sequences of SEQ ID
XX Nos 1 to 3032.

XX Sequence 1137 BP; 244 A; 250 C; 278 G; 365 T; 0 U; 0 Other;

XX Query Match 9.2%; Score 67; DB 14; Length 1137;
XX Best Local Similarity 59.9%; Pred. No. 1.1e-08;
XX Matches 112; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 473 CGCGGAGAAAGATGATGATCCCGGCAATATTTCAGTTAGATCTGTTCTAT 532
DB 773 CCGCGAAGAAAGACCGCGGAGCATCGAATATTCTAGTACGAGGGGTTCTGTGC 832
QY 533 TGGTGGCTACGCGATCTCTGTTAATCTTCGTTAGTTTCCGCGGTTACTCCGCG 592
DB 833 TCGTCTTCTCAGCATGATGATCTCATCTTCATTTGGTGGTCTCCCGCATTTCCCTCCG 892
QY 593 CTCGGTTTATGCGATTTGTTCCCATGCGGATATGTTTACTGCTGCTTGGCC 652
DB 893 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952
QY 653 TCATGCC 659
DB 953 TCATGCC 959

RESULT 10

XX ADX49328
XX ID ADX49328 standard; cDNA; 665 BP.

XX ADX49328;

XX 21-APR-2005 (first entry)

XX Plant full length insert polynucleotide seqid 24068.

XX plant protectant; plant growth regulator; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;

KM growth rate; cell cycle pathway; disease resistance;
 KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content; gene; ss.
 XX Unidentified.
 OS
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 PR 03-NOV-2001; 2001US-00985678.
 XX
 PA (LIU/J) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 24068; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SQ Sequence 665 BP; 111 A; 228 C; 205 G; 121 T; 0 U; 0 Other;
 Query Match 8.3%; Score 60.6; DB 13; Length 665;
 Best Local Similarity 64.7%; Pred. No. 6.4e-07;
 Matches 90; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

ADT18995
 ID ADT18995 standard; cDNA; 1055 BP.
 XX
 AC ADT18995;
 XX
 DT 13-JAN-2005 (first entry)
 XX
 DE Plant cDNA, Seq ID 4321.
 XX
 XX Plant; ss; gene; transgenic; cold tolerance; growth rate;
 KM drought tolerance; disease resistance; galactomannan production;
 KM plant growth regulator; heat tolerance; herbicide tolerance;
 KM lignin production; extreme osmotic condition tolerance;
 KM pathogens resistance; pest resistance; yield improvement; seed oil yield;
 KM seed protein yield.
 XX
 OS Viridiplantae.
 OS
 PN US2004216190-A1.
 XX
 PD 28-OCT-2004.
 XX
 PF 18-DEC-2003; 2003US-00739930.
 XX
 PR 28-APR-2003; 2003US-00424599.
 PR 28-APR-2003; 2003US-00425115.
 XX
 PA (KOVA/) KOVALIC D K.
 XX
 PI Kovalic DK;
 XX
 DR WPI; 2004-757369/74.
 XX
 PT New recombinant DNA constructs useful in the field of biotechnology and
 PT genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.
 XX
 PS Claim 1; SEQ ID NO 4321; 14pp; English.
 XX
 CC The invention relates a recombinant DNA construct comprising a
 CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
 CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
 CC Arabidopsis, wheat and rape but the specification does not indicate which
 CC sequences is derived from which organism. Also included is a method of
 CC producing a plant having an improved property, comprising transforming a
 CC plant with a recombinant DNA construct comprising a promoter region
 CC functional in a plant cell operably joined to a polynucleotide encoding a
 CC polypeptide associated with the property, and growing the transformed
 CC plant. The property is selected from improving plant cold tolerance, for
 CC manipulating growth rate in plant cells by modification of the cell cycle
 CC pathway, for improving plant drought tolerance, for providing increased
 CC resistance to plant disease, for galactomannan production, for production
 CC of plant growth regulators, for improving plant heat tolerance, for
 CC improving plant tolerance to herbicides, for increasing the rate of
 CC homologous recombination in plants, for lignin production, for improving
 CC plant tolerance to extreme osmotic conditions, for improving plant
 CC tolerance to pathogens or pests, for yield improvement by modification of
 CC photosynthesis, for modifying seed oil yield and/or content, for
 CC modifying seed protein yield and/or content, for yield improvement by
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
 CC and for yield improvement by providing improved plant growth and
 CC development under at least one stress condition. The polynucleotide may
 CC also encode a plant transcription factor. The methods and compositions of
 CC the present invention are useful in the field of biotechnology and
 CC genetics, in particular for producing transgenic plants with improved
 CC biological characteristics such as increased yield, improved nitrogen
 CC flow, increasing plant tolerance to cold or heat, improving plant
 CC tolerance to extreme osmotic and drought conditions, and improving plant
 CC tolerance to plant pests or pathogens. They can also be used in physical
 CC arrays of molecules, plant breeding markers, computer-based storage and
 CC analysis systems. The present sequence is one of the 5544 plant cDNA
 CC sequences of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC sequence.uspto.gov/sequence.html?docID=20040216190.
XX

Sequence 1055 BP; 218 A; 331 C; 268 G; 238 T; 0 U; 0 Other;

Query Match 8.1%; Score 59.6; DB 13; Length 1055;
Best Local Similarity 63.0%; Pred. No. 1.5e-06;
Matches 92; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 520 TCTCTGTCATGTTGGTGTCTAAGCATCTCTGTTAATACCTCCGTTAGTTTGGCG 579
DB 659 TCGTCTCTCTGCTCAATGTTGTCACCGCGTGGTGATTTCCCGCTGCTCCCA 718
QY 580 CCGTTACCTCCGCTCCGTTAATGCTGTTGTTCCCATTTGGGATTATGTTTACTG 639
DB 719 CCGCTGCTCCGCGCGCGCTGATGCTGATGTTGTCAGTGGCAATGCTGATGCTC 778
QY 640 GTGCTTCTTGCTTCATGCTCTTCTTC 665
DB 779 CTGGTCTGGCGTTTCATGCGACGTC 804

RESULT 12

ACT29210
ID ACL29210 standard; cDNA; 978 BP.

XX
AC ACL29210;

XX
DT 02-JUN-2005 (first entry)

XX
DE Rice abiotic stress responsive polynucleotide SEQ ID NO:3166.

XX
KM ss; abiotic stress tolerance; transgenic plant; plant; cereal;
XX agriculture.

XX
OS Oryza sativa.

XX
PN WO2003008540-A2.

XX
PD 30-JAN-2003.

XX
PF 21-JUN-2002; 2002WO-US019668.

XX
PR 22-JUN-2001; 2001US-0300112P.

XX
PR 24-AUG-2001; 2001US-0314662P.

XX
PR 26-SEP-2001; 2001US-0325277P.

XX
PR 21-NOV-2001; 2001US-0332132P.

XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX
PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
XX Moughamer T, Provart N, Rieke D, Zhu T;

XX
DR WPI; 2003-248011/24.

XX
PT New stress-responsive nucleic acid, useful for altering the
XX responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
XX stress, salt stress or osmotic stress.

XX
PS Claim 1; SEQ ID NO 3166; 89pp; English.

XX
CC The invention relates to novel abiotic stress responsive polynucleotides
XX and polypeptides. Also disclosed are vectors, expression cassettes, host
XX cells, and plants containing such polynucleotides. Also disclosed are
XX methods for using the polynucleotides and polypeptides to alter the
XX responsiveness of a plant to abiotic stress. The invention is useful in
XX agriculture. The nucleic acid is useful for determining whether a test
XX plant has been exposed to an abiotic stress condition. It is also useful
XX for selecting an agent that alters abiotic stress regulated
XX polynucleotide expression in a plant cell, and to identify a homolog or
XX ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
XX molecule and the polypeptide encoded by it are useful in altering the

CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
XX stress, osmotic stress or any of their combinations. The present sequence
XX is used in the exemplification of the invention

Sequence 978 BP; 253 A; 269 C; 210 G; 246 T; 0 U; 0 Other;

Query Match 7.7%; Score 56.6; DB 11; Length 978;
Best Local Similarity 59.7%; Pred. No. 1.1e-05;
Matches 95; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 505 TATTTCATGTTAAGTCTCTGTTCTATTGTTGTTAAGCATCTGTTAATACTT 564
DB 787 TACTTCAGCAGCAGGACCATCTGAGTGCATGCTCACCGTGTGCTGAGTCTG 846
QY 565 CCGTTAGTTTGGCGCGGTACCTCCGCTCCGTTAATGCTGATGTTCCATTGGG 624
DB 847 CCGCTCATCTGCGCGCGCTGCGCGCGCGCCGACAGCTGTGCTGCTGCGGTGTC 906
QY 625 ATTATGTTTACTCGTCTGTTCTTGCCCTTCATGCTCTT 663
DB 907 TTGCTGCGCTCTGTTGTTGCTGCTGCTTATGCCACT 945

RESULT 13

ABZ16156
ID ABZ16156 standard; DNA; 2000 BP.

XX
AC ABZ16156;

XX
DT 21-JAN-2003 (first entry)

XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 3961.

XX
KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX
OS Arabidopsis thaliana.

XX
PN WO200216655-A2.

XX
PD 28-FEB-2002.

XX
PF 24-AUG-2001; 2001WO-US026685.

XX
PR 24-AUG-2000; 2000US-0227866P.

XX
PR 26-JAN-2001; 2001US-0264647P.

XX
PR 22-JUN-2001; 2001US-0300111P.

XX
PA (SCRI) SCRIPPS RES INST.

XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.

XX
PT Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.

XX
PS Claim 144; SEQ ID NO 3961; 577pp + Sequence Listing; English.

XX
CC The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b) an
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
XX Sequence 2000 BP; 710 A; 365 C; 235 G; 690 T; 0 U; 0 Other;

Query Match 7.7%; Score 56; DB 6; Length 2000;
 Best Local Similarity 68.8%; Pred. No. 2.2e-05;
 Matches 77; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 270 CTGATCTTCTTAATCCAAACGACAGCAACCAACCTCTGTAGATGATTGAGAAAT 329
 |||||
 Db 1883 CTGATTTTCCAGATCAACTTCGAAAAGAGAAAAGCCCTTTTAATGATTCTGTAGTT 1942
 |||||

Qy 330 CTCGAACCTTCAAAAAGATATTATTAACATTCAAGACAGATTATTCGAACAC 381
 |||||
 Db 1943 CTCGAGCTCAAAAACGACATCATTAACATTCAAGACATTATCTCTCAAC 1994
 |||||

RESULT 14
 ADA69206
 ADA69206 standard; DNA; 2000 BP.
 AC ADA69206;
 XX
 XX 20-NOV-2003 (first entry)
 DT
 XX Arabidopsis thaliana gene, SEQ ID 2529.
 DE
 XX Plant; bacterial infection; fungal infection; viral infection; gene; ds.
 KM
 XX Arabidopsis thaliana.
 OS
 XX MO2003000898-A1.
 PN
 XX 03-JAN-2003.
 PD
 XX 22-JUN-2001; 2001MO-IB001105.
 PF
 XX 22-JUN-2001; 2001MO-IB001105.
 PR
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 DR WPI; 2003-175290/17.
 XX
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 PT
 XX Claim 27; SEQ ID NO 2529; 899bp; English.
 PS
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 CC
 SO Sequence 2000 BP; 710 A; 365 C; 235 G; 690 T; 0 U; 0 Other;

Query Match 7.7%; Score 56; DB 8; Length 2000;
 Best Local Similarity 68.8%; Pred. No. 2.2e-05;
 Matches 77; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 270 CTGATCTTCTTAATCCAAACGACAGCAACCAACCTCTGTAGATGATTGAGAAAT 329
 |||||
 Db 1883 CTGATTTTCCAGATCAACTTCGAAAAGAGAAAAGCCCTTTTAATGATTCTGTAGTT 1942
 |||||

Qy 330 CTCGAACCTTCAAAAAGATATTATTAACATTCAAGACAGATTATTCGAACAC 381
 |||||
 Db 1943 CTCGAGCTCAAAAACGACATCATTAACATTCAAGACATTATCTCTCAAC 1994
 |||||

Db 1943 CTCGAGCTCAAAAACGACATCATTAACATTCAAGACATTATCTCTCAAC 1994
 |||||

RESULT 15
 ID ADR64525
 ID ADR64525 standard; cDNA; 366 BP.
 AC ADR64525;
 XX
 XX 02-DEC-2004 (first entry)
 DT
 XX Cotton cDNA sequence, SEQ ID 5306.
 DE
 XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
 KM drought tolerance; plant disease resistance; galactomannan; lignin;
 KM plant growth regulator; heat tolerance; herbicide tolerance;
 KM homologous recombination; extreme osmotic condition tolerance;
 KM pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
 KM stress resistance.
 KM
 XX Gossypium hirsutum.
 OS
 XX US2004181830-A1.
 PN
 XX 16-SEP-2004.
 PD
 XX 29-JAN-2004; 2004US-00767795.
 PF
 XX 07-MAY-2001; 2001US-00849529.
 PR 12-DEC-2001; 2001US-00021323.
 PA (KONA/) KOVALIC D K.
 PA (ZHOU/) ZHOU Y.
 PA (CAOY/) CAO Y.
 XX
 XX Kovalic DK, Zhou Y, Cao Y;
 PI
 XX WPI; 2004-667718/65.
 DR
 XX New recombinant nucleic acid molecules and polypeptides from Gossypium
 PT hirsutum, useful for producing plants with improved biological
 PT characteristics (e.g. improved plant cold or drought tolerance).
 PT
 XX Claim 1; SEQ ID NO 5306; 14pp; English.
 PS
 XX The invention relates to a recombinant polynucleotide comprising any of
 CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
 CC Also a recombinant polypeptide comprising any of the 58798 amino acid
 CC sequences mentioned in the specification and producing a plant having an
 CC improved property. Producing a plant having an improved property
 CC comprises transforming a plant with a recombinant construct comprising a
 CC promoter region functional in a plant cell operably joined to a
 CC polynucleotide comprising a coding sequence for a polypeptide associated
 CC with the property, and growing the transformed plant. The polypeptide is
 CC useful for improving plant cold tolerance, manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, improving plant
 CC drought tolerance, providing increased resistance to plant disease,
 CC producing galactomannan (or lignin or plant growth regulators), improving
 CC plant heat tolerance, improving plant tolerance to herbicides, increasing
 CC the rate of homologous recombination in plants, improving plant tolerance
 CC to extreme osmotic conditions or to pathogens or pests, improving yield
 CC by modification of photosynthesis, modifying seed oil or protein yield
 CC and/or content, improving yield by modification of carbohydrate, nitrogen
 CC or phosphorus use and/or uptake, or improving yield by providing improved
 CC plant growth and development under at least one stress condition. The
 CC polynucleotide and polypeptide may also be used in recombinant DNA
 CC constructs, in physical arrays of molecules, as plant breeding markers,
 CC or in computer-based storage and analysis systems. The present sequence
 CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPRO at
 CC seqdata.uspro.gov/sequence.html?docid=20040181830. However only 6585
 CC polynucleotide sequences were available, the remaining 52213

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 03:56:47 ; Search time 185 Seconds
(Without alignments)
7033.379 Million cell updates/sec

Title:

US-10-715-129-1

Perfect score:

732

Sequence: 1 tttgtctccctccattcccta.....gaaatttatatgatgat 732

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCBUS_COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.6	6.8	7218	2	US-08-232-463-14
2	40.6	5.5	142783	3	US-09-949-016-15127
3	40	5.5	832	3	US-09-621-976-2813
4	39.8	5.4	1141	3	US-09-806-708B-22
5	38.4	5.2	2651	3	US-09-976-594-1050
6	38	5.2	197131	3	US-09-949-016-12675
7	38	5.2	197132	3	US-09-949-016-17170
8	37.6	5.1	64923	3	US-09-949-016-12901
9	37.6	5.1	65902	3	US-09-949-002-609
10	37.4	5.1	1141	3	US-09-806-708B-22
11	37.4	5.1	64593	3	US-09-949-016-16655
12	37.4	5.1	64593	3	US-09-949-016-16655
13	37.4	5.1	64593	3	US-09-949-016-16655
14	37.4	5.1	99498	3	US-09-949-016-12621
15	37.4	5.1	146428	3	US-09-949-016-12620
16	37.4	5.1	146438	3	US-09-949-016-12081
17	37.2	5.1	601	3	US-09-949-016-18101
18	37.2	5.1	601	3	US-09-949-016-18102
19	37.2	5.1	601	3	US-09-949-016-18103
20	37.2	5.1	601	3	US-09-949-016-51599
21	37.2	5.1	601	3	US-09-949-016-51600
22	37.2	5.1	601	3	US-09-949-016-51601
23	37.2	5.1	63658	3	US-09-949-016-13238
24	37.2	5.1	64489	3	US-09-949-016-11766

C 25	36.8	5.0	260286	3	US-09-949-016-17037	Sequence 17037, A
C 26	36.8	5.0	260293	3	US-09-949-016-12106	Sequence 12106, A
C 27	36.6	5.0	145287	3	US-09-949-016-13530	Sequence 13530, A
C 28	36.6	5.0	145287	3	US-09-949-016-13531	Sequence 13531, A
C 29	36.2	4.9	9821	3	US-08-956-171B-470	Sequence 470, App
C 30	36.2	4.9	9821	3	US-08-781-986A-470	Sequence 470, App
C 31	36.2	4.9	43095	3	US-09-676-519-17	Sequence 17, Appl
C 32	36	4.9	157644	3	US-09-949-016-16179	Sequence 16179, A
C 33	36	4.9	157644	3	US-09-949-016-16180	Sequence 16180, A
C 34	35	4.8	300	3	US-09-248-796A-12919	Sequence 15051, A
C 35	35	4.8	135667	3	US-09-949-016-15051	Sequence 12869, A
C 36	35	4.8	152486	3	US-09-949-016-12869	Sequence 12869, A
C 37	35	4.8	253345	3	US-09-949-016-12866	Sequence 12866, A
C 38	35	4.8	253364	3	US-09-949-016-13639	Sequence 13639, A
C 39	35	4.8	451924	3	US-09-949-016-12896	Sequence 12896, A
C 40	35	4.8	451925	3	US-09-949-016-17305	Sequence 17305, A
C 41	34.8	4.8	7218	2	US-08-232-463-14	Sequence 14, Appl
C 42	34.8	4.8	218940	3	US-09-949-016-17539	Sequence 17539, A
C 43	34.6	4.7	920	3	US-09-270-767-26471	Sequence 26471, A
C 44	34.6	4.7	1524	3	US-09-270-767-10971	Sequence 10971, A
C 45	34.6	4.7	2169	3	US-09-434-408-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHERFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-F18

US-08-232-463-14

Query Match
Best Local Similarity 6.8%; Score 49.6; DB 2; Length 7218;
Matches 19; Conservative 160; Mismatches 109; Indels 0; Gaps 0;

QY 1 TTGCTCTCATCTTCCCTACTAGTACTGTTTCACACAGTGTCTGATCCCAAC 60
D 1183 YY 1242
QY 61 CAATACCAAAAGCTTCAAACTCTTCACTCAAGCTTCTTCAATCG 120
D 1243 YY 1302
QY 121 TTGAGTTAACTCGATTTGTTCTGATCTGTTGGAATGCGGCACTCTATT 180
D 1303 YY 1362
QY 181 TGCTCGAATCTTCAACAAATGCTTGCATCAAGCTGATGTTAACAGTGCCTAA 240
D 1363 YY 1422
QY 241 AGATCAATCTTGAAGCAAAATTTGTCAGTCTTCAATCAAA 288
D 1423 YYYYYYYYYYGTACCAAAATCTTCTATCTCTTAATCACTTGCAATA 1470

RESULT 2

US-09-949-016-15127
Sequence 15127, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 15127
LENGTH: 142783
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(142783)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15127

Query Match
Best Local Similarity 5.5%; Score 40.6; DB 3; Length 142783;

Matches 85; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 272 GATCTCTTAATCCAAACAGACACAGCAAAACCACTCTGTAGATGTTGAGAAATCT 331
D 140872 GACCTTCCAAAAGCAAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 140931
QY 332 CAACCTTACAAAAGATTTTAAACATTCAAGACAGTTATTCGAACACCGAGTCAATG 391
D 140932 CACACACACAAAAGAGAAAACAACTAACAAATTTTAAAAAAGAACAAAATA 140991
QY 392 ACGTCGAGAAACAAACCGAAAAACATGAGCTTTCGA 430
D 140992 TAAGAGAAACAAAGACAAAACAAACGAAATGTGGA 141030

RESULT 3
US-09-621-976-2813

Sequence 2813, Application US/09621976
Patent No. 6639063

GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621, 976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 5.5%; Score 40; DB 3; Length 832;
Matches 43; Conservative 174; Mismatches 179; Indels 0; Gaps 0;

QY 178 TTTTGTCTGAATTTCTTCAACCAATTTGCTGATCAAGCTGATGTTAACAGAGTGGCC 237
D 5 WMYRTTYWYAKCWTMKMSWYMMYMKYMKYMKYMKYMKYMKYMKYMKYMKYMKYMKY 64
QY 238 TAAAGATCGATCTTGGAGCAAAATTTGTCAGTCTTCAATCCAAACAGACACA 297
D 65 WGYKKKAMCRITKTKKKKKGGYMMYTWGRBSYMAATTTWGYAYRSMYMKYMKYMKY 124
QY 298 GCAAAACAACTCTAGATGATTTGCAAGAAATCTCAAACTTACAAAAGATATTATTAAC 357
D 125 KKAAYRKTCYSSKCMTWKMKYMKYMKYMKYMKYMKYMKYMKYMKYMKYMKYMKYMKY 184
QY 358 ATTCAGACAGTTATTTGCAACCAACCGATCATGACCTCGAAGAAAACACCGAAGAAC 417
D 185 WMYKAKKSTWKRKSSYSASAKKCYSCSWAGAMKYMWMYMKYMKYMKYMKYMKYMKYMKY 244
QY 418 ATGAGCTTGAAGTTGCGCGGAGAAAAGCAAGAGTTAGGGCGAGTTCTCGCGG 477
D 245 ASCMRRKTAGSKSTBSKMMCTTSWKYCTKARTGTTCYRKGGMMGRGRMTASKK 304
QY 478 CAGAAAAGATGATGATCCCGGGAATTTATTCAGTTAGATGCTCTGTTCTATTGGTT 537
D 305 YMKRMMWMBABYRSTGTBRABMMYMYTWMMKMKYKAWABABMMYMMYMMYMMYMMY 364
QY 538 GGTCTAAGGATCTCTGTTAATATCTTCGTTAGTT 573
D 365 AAATATTAATTATTATGTAGCAATCTTGTACTTT 400

RESULT 4

US-09-806-708B-22
Sequence 22, Application US/09806708B

GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806, 708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patentin version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA

RESULT 7
US-09-949-016-17170/c

; Sequence 17170, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17170

; LENGTH: 197132

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(197132)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17170

Query Match 5.2%; Score 38; DB 3; Length 197132;
Best Local Similarity 54.2%; Pred. No. 3.6;
Matches 77; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 237 CTAAGATGATGATTTGACCAAAATTTGTCTACTGATCTTCTTAAACCAACGACAC 236

DB 44714 CTAAAGCCCAAGATTTTAAAGCAATAGTAGCTATGATGATGACGACGACGACCTTA 44655

QY 297 AGCAAAACAACCTCTGTAGATGATTCAGAAATCTCAACTTCAAAAGATATATATAA 356

DB 44654 GCGACACGACCAACAGCTGTCTTTAAATAATAATAATAATAATAATAATAAAG 44595

QY 357 CATTCAAGACGATTTTGAAC 378

DB 44594 GAATGAGAAAAGTTTATGATC 44573

RESULT 8
US-09-949-016-12901/c

; Sequence 12901, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12901

; LENGTH: 64923

; TYPE: DNA

; ORGANISM: Human

; GENERAL INFORMATION:

; APPLICANT: The University of British Columbia

US-09-949-002-609/c

; Sequence 609, Application US/09949002

; Patent No. 6900016

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: C1000790

; CURRENT APPLICATION NUMBER: US/09/949,002

; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/231,401

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 10823

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 609

; LENGTH: 65902

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(65902)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-002-609

Query Match 5.1%; Score 37.6; DB 3; Length 65902;
Best Local Similarity 52.6%; Pred. No. 2.6;
Matches 82; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 278 CTAATCCAAACGACACGACAAACCTGTAGATGATTCAGAAATCTCAACT 337

DB 40317 CTAAACCAACGACAAACCTGTAGATGATTCAGAAATCTCAACT 40258

QY 338 TACAAAGATATATTAACATTCAGACGATTTGCAACGACGATGACGCTG 357

DB 40257 AGAACAGATGTTATTAATAAACAAGAGCTCTGATTAATAAAGCATGACGTAG 40198

QY 398 GAAGAAACACCGAAGAAACATGACTTTCGAGTT 433

DB 40197 AAAAGTAAACTAAATAAACAGCTGGGTTGGAGAT 40162

RESULT 10
US-09-806-708B-22/c

; Sequence 22, Application US/09806708B

; Patent No. 6784342

; GENERAL INFORMATION:

; APPLICANT: The University of British Columbia

```

; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURES:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

Query Match          5.1%; Score 37.4; DB 3; Length 1141;
Best Local Similarity 11.4%; Pred. No. 0.3;
Matches 50; Conservative 150; Mismatches 237; Indels 0; Gaps 0;

QY 19 TACTAGTCTGTTTTCACACAGTTTCTGATCCACCAACCAACCAATACACAAAGCTTCTC 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 826 WABKHSWGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 79 AAATCTCTTCACTGCAAGCTTCTTCTTACATCTGTAATCTGTAAGTTAACTCGAGTTT 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 766 TTTTTCMCTKTMWNTWMDMTWMTBTYTTTNTSTMTNNNNNNNNNNNNNNNNNNNNNNNNNNNN
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 139 GTTCTGATCCTCTGTTTCTGAATCGTGGGACCTTATTTTGTCTGCAATCTTCACC 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 706 YTAATNTMCGMNNNTDARRNTNTTWRKMTNTKTRWSTTRKHHTGATNNNNNNNNNNNNNNNN
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 199 AATGCTTCATCAAGCTGATGTTTAAACAGTTGCCCTTAAAGATGATCTTTGAGCA 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 646 NNNNNNNNSCTCTRMWTTMTKMGDMTKKKYKMDTCTTYDVADSWVWVYAAWNR 587
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 259 AAATTTTGTACGATCTTTCTAAATCCAACGACACAGCAACCAACCACTCTGTATAGT 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 586 CRDVTYRNTTYCKSYAHSYWYWSNNAMWYRYSARWSSMARWTTNNNNNNNNNNNNNNNNNNNN
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 319 ATTCGAAATCTCAAACTTACAAAAGATATTATAACATTCAGACAGTATATTCGAAC 378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 526 AGTMMWNNNNNNNTDRTYNNWKKWABRTTYVDSMCNALSMMWGNWMMKMMWMAAND 467
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 379 AACCGAGTCATGACGCTCGGAAGAAACACGCAAAACATGAGCTTTGCAAGTTCCCG 438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 466 AGAMDHTTYMGNNTTMMRRABKGMMAKCRATCCNNNNNNRACVHKHKMRTWKYMW 407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 439 GAGAAAAGCAAGCAAGA 455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 406 KAAACNNNNBRAMWYRVA 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-09-949-016-16654/c
; Sequence 16654, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16654
; LENGTH: 64593
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)..(64593)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16654

Query Match          5.1%; Score 37.4; DB 3; Length 64593;
Best Local Similarity 51.5%; Pred. No. 3;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 550 TCTCTGTTAACTTCGCTAGTTGTTGCGCGCTTACCTCGCCCTCCGTTTAACTGCTA 609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14588 TCACTGAATTTACACAGTGTGTTTGTCTTCACTTCACTTCACTTCACTTCACTTCACTT
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 610 TTGGTTCCATTTGGGATTTATGTTTACTGCTGCTTCTGCTTCAAGCTTCTTCTCAT 669
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14528 TTACTTCATGCGCATCTGTTTGACGTTGAGTGCATCTTGTCTGTATTTTCTT 14469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 670 TCTAATGCTAATACAGATGTAATTCGCAATTTGATGTAATCTGAAA 716
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14468 ACAATGCCCAATTCGTCTTACCTGCTATATTGAATAATGGGATA 14422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
US-09-949-016-16655/c
; Sequence 16655, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16655
; LENGTH: 64593
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)..(64593)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16655

Query Match          5.1%; Score 37.4; DB 3; Length 64593;
Best Local Similarity 51.5%; Pred. No. 3;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 550 TCTCTGTTAACTTCGCTAGTTGTTGCGCGCTTACCTCGCCCTCCGTTTAACTGCTA 609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14588 TCACTGAATTTACACAGTGTGTTTGTCTTCACTTCACTTCACTTCACTTCACTTCACTT
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 610 TTGGTTCCATTTGGGATTTATGTTTACTGCTGCTTCTGCTTCAAGCTTCTTCTCAT 669
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14528 TTACTTCATGCGCATCTGTTTGACGTTGAGTGCATCTTGTCTGTATTTTCTT 14469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 670 TCTAATGCTAATACAGATGTAATTCGCAATTTGATGTAATCTGAAA 716
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14468 ACAATGCCCAATTCGTCTTACCTGCTATATTGAATAATGGGATA 14422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
RESULT 13
US-09-949-016-16656/c
; Sequence 16656, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16656
; LENGTH: 64593
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(64593)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16656

Query Match
Best Local Similarity 51.5%; Pred. No. 3;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 550 TCTCTGTTAATCTTCGCTAGTTGTCGCCGCTTACCTCCGCTCGTTATGCTGTA 609
Db 14588 TCACGTAAATTACACAGTGTGTTTGTCTTCACTTGACCTCCTAAGTTATGTA 14529

Qy 610 TTGGTCCCATGGAGATTAAGTTTACTGCTGCTTCCCTTCATGCTTCTTCAT 669
Db 14528 TTACTTCCATGGGCACTGTTGTGACGTTGAGTTGCATTTGCTTGTATTTCTT 14469

Qy 670 TCTAATGCTAATACAGATGTAATTCGCAATTTCAATTAATCTGAAA 716
Db 14468 ACAATGCCCAATTCGTCTTACCTGTCTAATTAATTAATGAGATA 14422

RESULT 14
US-09-949-016-12621/c
; Sequence 12621, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12621
; LENGTH: 99498
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(99498)
```

```
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12621

Query Match
Best Local Similarity 51.5%; Pred. No. 3.8;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 550 TCTCTGTTAATCTTCGCTAGTTTGGCCGCTTACCTCCGCTCGTTATGCTGTA 609
Db 49493 TCACGTAAATTACACAGTGTGTTTGTCTTCACTTGACCTCCTAAGTTATGTA 49434

Qy 610 TTGGTCCCATGGAGATTAAGTTTACTGCTGCTTCCCTTCATGCTTCTTCAT 669
Db 49433 TTACTTCCATGGGCACTGTTGTGACGTTGAGTTGCATTTGCTTGTATTTCTT 49374

Qy 670 TCTAATGCTAATACAGATGTAATTCGCAATTTCAATTAATCTGAAA 716
Db 49373 ACAATGCCCAATTCGTCTTACCTGTCTAATTAATTAATGAGATA 49327

RESULT 15
US-09-949-016-12620/c
; Sequence 12620, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12620
; LENGTH: 146428
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(146428)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12620

Query Match
Best Local Similarity 51.5%; Pred. No. 4.7;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 550 TCTCTGTTAATCTTCGCTAGTTTGGCCGCTTACCTCCGCTCGTTATGCTGTA 609
Db 48423 TCACGTAAATTACACAGTGTGTTTGTCTTCACTTGACCTCCTAAGTTATGTA 48364

Qy 610 TTGGTCCCATGGAGATTAAGTTTACTGCTGCTTCCCTTCATGCTTCTTCAT 669
Db 48363 TTACTTCCATGGGCACTGTTGTGACGTTGAGTTGCATTTGCTTGTATTTCTT 48304

Qy 670 TCTAATGCTAATACAGATGTAATTCGCAATTTCAATTAATCTGAAA 716
Db 48303 ACAATGCCCAATTCGTCTTACCTGTCTAATTAATTAATGAGATA 48257
```

Search completed: January 30, 2006, 06:39:09
Job time : 189 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: January 30, 2006, 08:30:29 ; Search time 2493 Seconds
(without alignments) 2416.930 Million cell updates/sec

Title: US-10-715-129-2

Perfect score: 533
Sequence: 1 MDVGRNNRKMSPRSSPEKS.....LAFWPSHSNANTDVTCNFM 106

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlp
-Q/cgnt1/USPRO.spool.p/US10715129/runat.27012006.154744.14566/app.query.fasta.1.263
-DB=genembl -QPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LCOPLC=0 -LCOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10715129 -QCGN 1 1 7415 -Qrunat.27012006.154744.14566 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: GenEmbl.*

- 1: gb Da:*
- 2: gb In:*
- 3: gb Env:*
- 4: gb Om:*
- 5: gb Ov:*
- 6: gb Pat:*
- 7: gb Ph:*
- 8: gb Pr:*
- 9: gb Ro:*
- 10: gb Sts:*
- 11: gb Sy:*
- 12: gb Un:*
- 13: gb Vi:*
- 14: gb Htg:*
- 15: gb Pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Length	DB ID	Description
1	533	100.0	393	BT011724 Arabidops
2	533	100.0	732	AY305869 Arabidops
3	533	100.0	747	AK176640 Arabidops

SUMMARIES

4	533	100.0	100906	15	ATP24G16	AL138647 Arabidops
5	327	61.4	336	6	AX506572	AX506572 Sequence
6	327	61.4	336	6	AX651235	AX651235 Sequence
7	327	61.4	439	15	BT000930	BT000930 Arabidops
8	327	61.4	904	15	AY080817	AY080817 Arabidops
9	327	61.4	109741	15	AC004005	AC004005 Arabidops
10	224.5	42.1	1027	15	AK061481	AK061481 Oryza sat
11	224.5	42.1	1137	15	AK060546	AK060546 Oryza sat
12	224.5	42.1	110000	15	AP008217_078	Continuation (79 O
13	224.5	42.1	180999	15	AC137924	AC137924 Oryza sat
14	209	39.2	1302	15	AK102641	AK102641 Oryza sat
15	209	39.2	46422	15	OSJN00249	AL731506 Oryza sat
16	209	39.2	90825	15	AC157500	AC157500 Oryza sat
17	209	39.2	110000	15	AP008210_221	Continuation (222
18	209	39.2	118959	15	OSJN00235	Continuation (79 O
19	197.5	37.1	110000	15	AP008218_057	Continuation (58 O
20	197.5	37.1	159056	15	CNS080C7	AL954853 Oryza sat
21	179.5	33.7	117439	15	AC122163	AC122163 Medicago
22	179.5	33.7	154378	15	AC126006	AC126006 Medicago
23	177	33.2	767	15	AK070163	AK070163 Oryza sat
24	177	33.2	110000	15	AP008207_386	Continuation (387
25	177	33.2	191022	15	AP003431	AP003431 Oryza sat
26	170.5	32.0	629	15	DQ090057	DQ090057 Archichs s
27	167.5	31.4	77319	15	AP004474	AP004474 Lotus cor
28	167.5	31.4	91654	15	AP006393	AP006393 Lotus cor
29	164.5	30.9	3352	10	BV144464	BV144464 PZA02432-
30	163.5	30.7	107407	15	AC134933	AC134933 Oryza sat
31	163.5	30.7	110000	15	AP008211_203	Continuation (204
32	157	29.5	559	15	AK070327	AK070327 Oryza sat
33	157	29.5	1044	15	AK061571	AK061571 Oryza sat
34	157	29.5	110000	15	AP008214_007	Continuation (8 of
35	157	29.5	185545	15	AP005657	AP005657 Oryza sat
36	153.5	28.8	330	10	BV146385	BV146385 PZA02609-
37	153.5	28.8	343	10	BV146386	BV146386 PZA02609-
38	153.5	28.8	345	10	BV146389	BV146389 PZA02609-
39	151	28.3	111256	15	AC125478	AC125478 Medicago
40	149.5	28.0	342	10	BV146383	BV146383 PZA02609-
41	149.5	28.0	342	10	BV146384	BV146384 PZA02609-
42	149.5	28.0	342	10	BV146388	BV146388 PZA02609-
43	147.5	27.7	342	10	BV146393	BV146393 PZA02609-
44	145	27.2	32163	14	AC160455	AC160455 Glycine m
45	144.5	27.1	326	10	BV146392	BV146392 PZA02609-

ALIGNMENTS

RESULT 1	BT011724	393 bp	mRNA	linear	PLN 04-MAR-2004
LOCUS	BT011724				
DEFINITION	Arabidopsis thaliana At3g59900 gene, complete cdb.				
ACCESSION	BT011724				
VERSION	BT011724.1	GI:44917524			
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Chen, R., Chen, H., Kim, C. J., Shinn, P., Carninci, P., Hayashizaki, Y., Itohida, J., Kamiya, A., Kawai, J., Narusaka, M., Sakurai, T., Satou, M., Seki, M., Shinozaki, K., and Ecker, J. R.				
AUTHORS	Arabidopsis ORF clones				
TITLE	Unpublished				
JOURNAL	Submitted (04-MAR-2004) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA				

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGRS (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Shinn, P., Bowser, L., Chan, M. M., Chang, C. M., Dale, J. M., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PI.

FEATURES

source

Location/Qualifiers

1. .393

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="U65994"

/ecotype="Columbia"

/note="unknown protein"

/product="At3g5990"

/protein_id="AA049087.1"

/db_xref="GI:44917525"

/translation="MIRISNLODIIINODSYNNRVMDVGNRRKNSFRSPSKS KOELRSFSAOKMMIPANFYSLSLFLVGLVLAASLLILPLVLPPLPPFMLLVPI

GIWLVLAIFMPSHSNANTDVTGCM"

ORIGIN

Alignment Scores:

Pred. No.: 1,11e-45 Length: 393
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-715-129-2 (1-106) x BT011724 (1-393)

QY 1 MetAspValGlyArgAsnAsnArgLysAsnMetSerPheArgSerProGluLysSer 20
Db 73 ATGACGTCGGAGAAACAAACCGGAAACATGAGCTTTCGAAAGTTCGCGGAGAAAGC 132
QY 21 LysGlnGluLeuArgArgSerPheSerAlaGlnLysArgMetMetIleProAlaAsnTyr 40
Db 133 AAGCAAGATTACGCGGAGTTTCTCGCGCAAAAGATGATGATCCGCGCAATTAT 192
QY 41 PheSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuIleLeuPro 60
Db 193 TTCAGTTAAGCTCTGTTCCATGCTGTTGCTTAACGCGCATCTCTGTTAATACCTCCG 252
QY 61 LeuValLeuProProLeuProProProPheMetLeuLeuLeuValProIleGlyIle 80
Db 253 TTAGTTTGGCGCGGTTACCTCCGCTCCGCTTAATGCTGCTATGTTCCATTCGGAATT 312
QY 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsnThrAsp 100
Db 313 ATGGTTTACTCGCGCTTCTTGCTTCAATGCTTCTTCTCAATTCCTAATGCTAATACAGAT 372
QY 101 ValThrCysAsnPheMet 106
Db 373 GTAACCTGCAATTCATG 390

RESULT 2
AY305869

AY305869 732 bp mRNA linear PLN 03-SEP-2003

DEFINITION

Arabidopsis thaliana auxin-inducible protein (ARGOS) mRNA, complete cds.

ACCESSION
AY305869
VERSION
AY305869.1
KEYWORDS
GI:34329825

SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLES

JOURNAL

PUBMED

Qy 101 ValThrcCysaenphenet 106
Db 688 GTAACCTGCATTTTCATG 705
RESULT 3
AKI16640
LOCUS
DEFINITION Arabidopsis thaliana mRNA for putative protein, complete cds,
 clone: RAFL25-17-J07.
ACCESSION AKI16640
VERSION AKI16640.1
KEYWORDS GI:51971476
SOURCE FLI CDNA.
ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
AUTHORS Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,
 Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K.,
 Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,
 Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
 Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs
 Unpublished
 2 (bases 1 to 747)
 Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,
 Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K.,
 Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,
 Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
 Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Direct Submission
 Submitted (06-SEP-2004) Motoaki Seki, RIKEN Genomic Sciences
 Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
 230-0045, Japan (E-mail: msekik@sec.riken.jp,
 URL: http://large.gsc.riken.jp/, Tel: 81-45-503-9625,
 Fax: 81-45-503-9586)
COMMENT An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al. (1998) Plant J. 15:707-720,
 Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
 and XhoI was ligated to modified lambdaPhiC-1 vector (Carninci et
 al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This
 clone is in a modified pBluescript vector.
 Please visit our web site (http://large.gsc.riken.jp/) for further
 details.
FEATURES
source
 1. 747
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAFL25-17-J07"
 /ecotype="Columbia"
 /note="common name: thale cress"
 1. 747
 /gene="At3g59900"
 314. 706
 /gene="At3g59900"
 /codon_start=1
 /product="putative protein"
 /protein_id="BAD4403.1"
 /db_xref="GI:51971477"
 /translation="MIRISLQLQDIINIDSYNNRVMVDGRRNRKMSRRSSPEKS
 KQIRLRSFSAQKRMWIPANYSLSLFLVGLTASLILPLVPLPPPPMLLVPI
 GIVLVLVLAFMPSHSNANTDVICNMF"
ORIGIN
 Alignment Scores:
 Pred. No.: 2.17e-45 Length: 747
 Score: 533.00 Matches: 106
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0

DB: 15 Gaps: 0
US-10-715-129-2 (1-106) x AKI16640 (1-747)
Qy 1 MetApValGlyArgAenAenArgLyAaMwSerPheArgSerProGlyLysSer 20
Db 386 ATGACCGTCGGAGAAACCAACCGGAAAAACATGAGCTTTCGAGAGTCCGCGAGAAAGC 445
Qy 21 LygGlnGluLeuAgaGserPheSerAlaGlnlyArgMetMetIleProAlaAsnTyr 40
Db 446 AAGCAAGATTACCGCGAGATTCTCGCGCGAAGAAAGAAAGATGATCCCGCAATTAT 505
Qy 41 PheSerLeuGlnSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuIleLeuPro 60
Db 506 TTCGTTTAGAGTCTCTGTTCTATGAGTGTCTAACGGCATCTCTGTAACTTCG 565
Qy 61 LeuValLeuProProLeuProProProPheMetLeuLeuValProIleGlyIle 80
Db 566 TTGATTTTGGCGCGCTTACCTCGCGCTCCGTTTATGCTGCTATGTTCCCATGGAGTT 625
Qy 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAlaAsnThrAsp 100
Db 626 ATGGTTTACTCGTCGCTTCTTGCCCTTCATGCTCTCTCAATGCTAAATACAGAT 685
Qy 101 ValThrcCysaenphenet 106
Db 686 GTAACCTGCATTTTCATG 703
RESULT 4
ATP24G16
LOCUS
DEFINITION Arabidopsis thaliana DNA chromosome 3, BAC clone P24G16.
ACCESSION AL138647
VERSION AL138647.1
KEYWORDS GI:6899904
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
AUTHORS D'Angelo, M., Vezzi, A., Modesto, D., Pigazzi, M., Valle, G.,
 Mewes, H.W., Lemcke, K., Mayer, K.F.X., Quetier, F. and Salanoubat, M.
 Unpublished
 2 (bases 1 to 100906)
JOURNAL Arabidopsis sequencing project.
REFERENCE EU Arabidopsis sequencing project.
AUTHORS Direct Submission
TITLE Submitted (18-FEB-2000) MIPS, at the Max-Planck-Institut fuer
 Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
 lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project
 Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
 d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue
 Gaston Cremieux, BP191, 91006 Evry Cedex, France;
 http://www.genoscope.cns.fr
COMMENT Information on performance of analysis and a more detailed
 annotation of this entry and other sequences of chromosomes 3, 4
 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
FEATURES
source
 1. 100906
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /ecotype="Columbia"
 1. 1731
 /note="overlap to BAC F16L24, please refer EMBL Acc. no
 AL138659 for analysis and annotation"
 2. 1888
 /gene="F24G16.10"
 /complement (<2. 1888)
 /gene="F24G16.10"
 /complement (<2. 1888)
 /gene="F24G16.10"

exon complement (7549, .7658)
/gene="F24G16.30"
/number=9
intron complement (7659, .7818)
/gene="F24G16.30"
/number=9
exon complement (7819, .8175)
/gene="F24G16.30"
/number=10
gene 12111, .18504
/gene="F24G16.40"

Alignment Scores:

Pred. No.: 3,49e-43 Length: 100906
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-715-129-2 (1-106) x ATF24G16 (1-100906)

Qy 1 MetAspVal1G1YArGAsnAsnArgLyAsnMetSerPheArgSerSerProGlnLysSer 20
Db 62943 ATGACCGTGGAGAAACACCGGAAAACATGAGCTTCCAGATCCCGGAGAAAAGC 63002
Qy 21 LysGlnGlnLeuArgArgSerPheSerAlaGlnLysArgMetMetLleProAlaAsnTyr 40
Db 63003 AAGCAAGATTACGGCGGAGTTTCTCGCGCAAAAAGATGATGATCCCGCGAATTAT 63062
Qy 41 PheSerLeuGlnSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuLleLeuPro 60
Db 63063 TTCAAGTTTAGAGTCTCTGTTCTTATGGTGGCTTAAGGCAATCTCTGTATATCTTCG 63122
Qy 61 LeuValLeuProProLeuProProProPheMetLeuLeuValProLleGlylle 80
Db 63123 TTAAGTTTGGCCGCGCTTACCTCCGCTTATAGCTGCTATTTGTTCCCATTTGGAGATT 63182
Qy 81 MetValLeuLeuVal1ValLeuAlaPheMetProSerSerAlaSerAsnAlaAsnThrAsp 100
Db 63183 ATGGTTTATCGTCGTCCTTCTTGCTTCAATGCTTCTTCAATTCATTAAGCTTAATCAAT 63242
Qy 101 ValThrCyAsnPhenMet 106
Db 63243 GTAACCTGCATTTTCATG 63260

RESULT 5
AX506572 LOCUS AX506572 336 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1267 from Patent WO0216655.
ACCESSION AX506572
VERSION AX506572.1 GI:23387809

KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1267 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)

FEATURES
source location/Qualifiers
1..336
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:

Pred. No.: 1.3e-24 Length: 336
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: 6 Gaps: 3

US-10-715-129-2 (1-106) x AX506572 (1-336)

Qy 1 MetAspVal1G1YArGAsnAsnArgLyAsnMetSerPheArgSerSerProGlnLysSer 16
Db 1 ATGACCGTGGAGAAACACCGGAAAACATGAGCTTCCAGATCCCGGAGAAAAGC 60
Qy 17 ProGlnLysSerLyGlnGlnLeuArgArgSerPheSerAlaGlnLys-----ArgMet 34
Db 61 CCGATTATGGGAGCAAGAAATTTTTCGACATTTGCTGTCGCAACAGTCCCAAGAGG 120
Qy 35 MetLleProAlaAsnTyrPheSerLeuGlnSerLeuPheLeuValGlyLeuThrAla 54
Db 121 CTAAATACAGGAGATTACCTTCAAGTTTAGAATCAATGCTTGCTTGGTCTCACAGCA 180
Qy 55 SerLeuLeuLleLeuProLeuValLeuProProProProPheMetLeuLeu 74
Db 181 TCTCTTGAATCTTACCGTTGATTTCTTCCACATTTGCTCTCTCTTTATGCTGCTT 240
Qy 75 LeuValProLleGlylleMetValLeuLeuVal1ValLeuAlaPheMetProSerSerHis 94
Db 241 TTGATTCCTAATGGGATTAATGTTTGGTATGTTGCTTCAATGCTTCTTCAAT 300
Qy 95 Ser 95
Db 301 TCC 303

RESULT 6
AX651239 LOCUS AX651239 336 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 23 from Patent WO03000898.
ACCESSION AX651239
VERSION AX651239.1 GI:29154057

KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Karagiannis, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 23 03-JAN-2003;
Syngenta Participations AG (CH)

FEATURES
source location/Qualifiers
1..336
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:

Pred. No.: 1.3e-24 Length: 336
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: 6 Gaps: 3

US-10-715-129-2 (1-106) x AX651239 (1-336)

Qy 1 MetAspVal1G1YArGAsnAsnArgLyAsnMetSerPheArgSerSerProGlnLysSer 16
Db 1 ATGACCGTGGAGAAACACCGGAAAACATGAGCTTCCAGATCCCGGAGAAAAGC 60
Qy 17 ProGlnLysSerLyGlnGlnLeuArgArgSerPheSerAlaGlnLys-----ArgMet 34

```

Db      61  CCGATTATGGGGAAGCAAGATTGTTTCGACATGTCGACAGAACATGCCAAGGAGG 120
Qy      35  MetIleProIlaaenTyRphSerLeuGluSerLeuPheLeuValGlyLeuThAla 54
Db      121  CTAATATCAGGAGATCTTCACTTCAATCAATGATGTTGCTGTTGCTTCAACAGCA 180
Qy      55  SerLeuLeuLeuProLeuValLeuProProLeuProProProProMetLeuLeu 74
Db      181  TCTCTTGAATCTTACGCTGATCTTCCACCATTCCTCTCTCTTATATCTCTT 240
Qy      75  LeuValProIleGlyIleMetValLeuValLeuValLeuAlaPheMetProSerSerHis 94
Db      241  TTGATTCCTATGGGATATGATGTTTGCTATGCTTCTGCTTCAATGCTTCTTCAAT 300
Qy      95  Ser 95
Db      301  TCC 303

RESULT 7
BT000930
LOCUS   Arabidopsis thaliana clone C105228 unknown protein (At2g44080)
DEFINITION
ACCESSION   BT000930.1 GI:24030324
VERSION     BT000930
KEYWORDS    FLI CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana

REFERENCE
AUTHORS     Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
            Nguyen,M., Palm,C.J., Shim,P., Southwick,A., Tripp,M.G., Wu,T.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
            Arabidopsis Open Reading Frame (ORF) Clones
            Unpublished
            2 (bases 1 to 439)

TITLE       Arabidopsis thaliana (thale cress)
JOURNAL     Submitted (16-OCT-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     Annotation based on January 2002 version of the Arabidopsis genome
            submitted to Genbank.

FEATURES
SOURCE
LOCATION/Qualifiers
1..439
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="C105228"
/ecotype="Columbia"
/notes="This clone is in pUNI 51."
1..439
/gene="At2g44080"
1..408
/gene="At2g44080"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA041330.1"
/db_xref="GI:24030325"
/translation="MIREPSLDIINIOEHSINNMVGRGHRKATSPFGSAPA
            PINKQELPRTLSQNSPRRLISATYSLSMVVLVLTSLTLPILPLPPPPPM
            LTLIPGIVLWLVLAIPSSNSKRVSSSTFM"
409..439
3' UTR

```

```

ORIGIN
/gene="At2g44080"

Alignment Scores:
Pred. No.: 1,72e-24 Length: 439
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: 15 Gaps: 3

US-10-715-129-2 (1-106) x BT000930 (1-439)
Qy      1  MetAspVal---GlyArgaenAArgLysaMetSerPheArgSer----- 16
Db      73  ATGACCTGAGAGAGATCAATACCGAATAACAGAGTTTCTGCTTCAAGCTCAGCT 132
Qy      17  ProGluysSerLysGlnGluLeuArgSerPheSerAlaGlnys-----ArgMet 34
Db      133  CCGATTATGGGGAAGCAAGATTGTTTCGACATTCGTCGCGAAGACATGCCAAGAGG 192
Qy      35  MetIleProIlaaenTyRphSerLeuGluSerLeuPheLeuValGlyLeuThAla 54
Db      193  CTAATATCAGGAGATCTTCACTTCAATCAATGATGTTGCTGTTGCTTCAACAGCA 252
Qy      55  SerLeuLeuLeuProLeuValLeuProProLeuProProProProMetLeuLeu 74
Db      253  TCTCTTGAATCTTACGCTGATCTTCCACCATTCCTCTCTCTTATATGCTCTT 312
Qy      75  LeuValProIleGlyIleMetValLeuValLeuValLeuAlaPheMetProSerSerHis 94
Db      313  TTGATTCCTATGGGATATGATGTTTGCTATGCTTCTGCTTCAATGCTTCTTCAAT 372
Qy      95  Ser 95
Db      373  TCC 375

RESULT 8
AY080817
LOCUS   Arabidopsis thaliana unknown protein (At2g44080) mRNA, complete
DEFINITION
ACCESSION   AY080817 GI:19423989
VERSION     AY080817.1
KEYWORDS    FLI CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana

REFERENCE
AUTHORS     Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
            Setou,M., Seki,M., Shim,P., Southwick,A., Shinzaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
            Arabidopsis Full Length cDNA Clones
            Unpublished
            2 (bases 1 to 904)

TITLE       Arabidopsis thaliana (thale cress)
JOURNAL     Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the

```

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Saitou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shimozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banb, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldenfild, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shimozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source
location/Qualifiers
1..904
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAF06-84-F13 (R11911)"
/ecotype="Columbia"
/note="This clone is in a modified pBluescript vector (Lambda ZAP) as a XhoI/SacI insert."
1..904
/gene="At2g44080"
1..443
/gene="At2g44080"
444..851
/gene="At2g44080"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA087236.1"
/db_xref="GI:19423990"
/translation="MIREFSSLDNDINIOEHYSLNNMDVGDHNRKNTSPRSADA
PIHGKOLPRTLSQNSPRLISASYSLSMVLVGLTASLLPLPLPPPPM
LLPLPISIMVLMLAPMPSPNSKRVSSSTFM"
852..904
/gene="At2g44080"

ORIGIN

3' UTR

Alignment Scores:
Pred. No.: 3,636-24 Length: 904
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
15 Gaps: 3
US-10-715-129-2 (1-106) x AY080817 (1-904)

1 MetAspVal---GlyArgAsnAsnArgGlyAsnMetSerPheArgSerSer----- 16
516 ATGACCTGAGAGAGATCATACCGAAGAAACAGAGCTTTCTGCTGCTCCAGCT 575
17 ProGluyseryLgIngluLeuArgSerPheSerAlaGlnLys-----ArgMet 34
576 CCGATTATGGGGAAGCAAGATTGTTCCGACATTGTCGTGCGAAGACAGTCCAGAGG 635
35 MetIleProAlaAenTyrPheSerLeuGlnSerLeuPheLeuValGlyLeuThrAla 54
636 CTAATATCAAGGAGTACTTCAAGTTTGAATCAATGATGTGCTGTGTTGCTCAAGCA 695
55 SerLeuLeuIleuLeuProLeuValLeuProProLeuProProProMetLeuLeu 74
696 TCTCTTGAATCTTACCGTATGATCTTCCACCATTTGCTCTCTCTTATATGCTGTT 755

Qy 75 LeuValProIleGlyIleMetValLeuValLeuValLeuAlaPheMetProSerSerHis 94
Db 756 TTAGTCTTATGGGATATGATGTTTCTATATGATGTTCTTCTTCAATGCTCTTAAAT 815

Qy 95 Ser 95
Db 816 TCC 818

RESULT 9

AC004005

LOCUS

DEFINITION

AC004005

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

TITLE

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

misc_feature

repeat_region

gene

mRNA

CDS

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

gene
7192..8658
/gene="At2g43890"
/note="synonym: F6B13.2"
join(<7192..7632,7743..8032,8122..8350,8440..>8658)
/gene="At2g43890"
join(7192..7632,7743..8032,8122..8350,8440..8658)
/gene="At2g43890"
/product="putative polygalacturonase"
/protein_id="AAC23398.1"
/db_xref="GI:3212847"
/translation="MDNNKLVAVLLMFFSSPLMKSTAASTNVVSGAKPDGRTDS
TKAFGLMWAACRAAAVTVTPVSGFLKPVFSGFRIRITFOIYGLITVADSVGS
LNGSWILFQAKNRIISIGTLDAKASFMACKRSKSGASMTFVADVAVS
GLTISGOTTHLVINSQNVIVRKVLVAPDOSPTNGDVLQSGAGTVLDGPHNGD
DCISIGOTNRLVMSKLCNCPGHCISIGSGRANPGVNTLINSVPSGSDNGRI
KTVAROSTGFVNVLFQNLTKNVONFTIVDNYCPBNQCPKQSGVKLSIVVYNI
QGSTRTOALTFDCSRSPCOAIRLHDIKLTFNGRSATSTCKNIKGVKGVNVPQCL
"
CDS
complement(9101..14090)
/gene="At2g43900"
/note="synonym: F6B13.3"
complement(join(<9101..9952,10049..10409,10500..10623,
10717..10984,11093..11133,11219..11364,11587..11671,
11768..11895,12004..12811,12910..13162,13259..>14090))
/gene="At2g43900"
complement(join(9101..9952,10049..10409,10500..10623,
10717..10984,11093..11133,11219..11364,11587..11671,
11768..11895,12004..12811,12910..13162,13259..14090))
/gene="At2g43900"
/codon_start=1
/product="putative inositol polyphosphate 5'-phosphatase"
/protein_id="AAC23399.1"
/db_xref="GI:3212848"
/translation="MDTINNHRDENDDEEALSAMSVPPRKTHSYHQIARQ
KGRHROHSLDIPKITEIVSGGIGSDSDDEFPYATTTSSSEFPFGDGTGSD
DYLOPEIGEDFQPLPEFVSGGAGVGFKPTPSPLHSARPCILEPFLKSTQVR
FLINACTETOLMAGSGSVRFNNFDDAFERGGLSGRVORGEDAAPFESASTPT
TCMVNNGNRLVNSGHDKGIRSMKNVYVLDGDDSPFKGSLMOAKHGQNVNCS
YGVLMCSGCGVYKIKTWBSMEKSLRLTEKEMALLVERSGIDLRQVTTNGTCS
SSSEVKCLADNVRSKVMAOLOTFSIMDRTELELVNVEGQTERNVMPJGQDP
LAADDMKAKIATGTSKKEKPGFLORSNNAAGADAVRAVRVGGAGVBERKTEMT
LADGIMWTGCTGGLIQMDGNGRLDDFRRHOCAYCFCTPERIYIGVSGHIOI
DLEGNLLAGVANNNAVIRMAADGYFSLATGSGIRGMPVISRGLDGIIRLEBLEK
ERTYAOITDSVRLITGSMNVGOGKASHDALMSYGSVADYGLIYVGLQREMGAGLA
MSAKSVSGNESTTICQYWTDTIGTLDKKAVERKSGRLAGLISLWVKRLKTH
VGIIDVAVPCGFRALGNKGVGGLRIYRVDRLMCFNLHAHLAEVNRNADPHI
YKMSFTSRSNANHPAGVSTGSHTKSNANNAVNTRETKODLAEDAVVFGDFPY
RLFGISYDEARDVFSQSFMLREKQDLRAEMKARGVFCQMRBAITFPPTFEERH
PGAGYDSGEKKRIIPACDRVIFPDRTSPESGSLDPCPVASIMLYDADCMQVTESSH
KVPKRCFKVKEIHDVRSVROBFGRITKTEKVRALLDLRYVPTIVSNSIVLQNO
DITVLRITTKCKENAVFRLCBQSTVRDEDTLEAPLISGSGFPMLEVMPPAGIT
KIPSSVEVSVHHEFHTLEFVDGIQNNWCEBTRDEALLVAVNOGSGCTEVNCHRV
HVHCFSAKNRIIDSNPSNKSQSLKKNBDSNKSXKSDGSDNSKSKSGSDGS
SKSKSPKSDGSDNSKSKSDGSDNSKSKSDGSDNSKSKSDGSDNSKSKSDGS
CSXSKSGSDGSDNSKSKSDGSDNSKSKKNDGSDSKSKKNDGSDSSKSKKSG
DSSSKSHKSKSGSDSLTRRTMEILLQHTSRVGRKTVTRQL"
repeat_region
14103..14213
/rpt_family="(GAGAA)n"
/rpt_protein_id="AAC23400.1"
/note="synonym: F6B13.4; supported by full length cDNA:
Ceres.24370"
complement(join(<14748..14963,15285..15363,15476..15522,
15617..15716,15874..15974,16059..16145,16677..16807,
17119..>17292))
/gene="At2g43910"
complement(join(14925..14963,15285..15363,15476..15522,
15617..15716,15874..15974,16059..16145,16677..16807,
17119..17218))
/gene="At2g43910"
/codon_start=1

gene
23418..25849
/gene="At2g43930"
/note="synonym: F6B13.6; contains a protein kinase domain
profile (PDOC00100)"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
/codon_start=1
/product="putative protein kinase"
/protein_id="AAC23427.1"
/db_xref="GI:3212876"
/translation="MDHRRRLATIKILIRNSSTLLFPKIVPILRLSGGDSRALPYT
PFQCSRALNIGHNTSRSEPHVILQCKRKSVCPSSEFHLCLFGYIDEDERKHF
GHKGVVAREEROLLKNSHPNIVYIDGTVRSDDLNTMMQVPGGSLSLAKXGS
FPEPKHNRVLDNAISHTSYKNIHQGSFKLEIDPRKKGS"
complement(27269..30414)
/gene="At2g43940"
/note="synonym: F6B13.7"
complement(join(<27269..27361,27641..27771,27853..27984,
28336..28397,28479..28537,28635..28784,28865..28981,
29149..29343,30043..>30414))
/gene="At2g43940"
repeat_region
complement(15108..15141)
/rpt_family="AT_rich"
complement(16570..16608)
/rpt_family="AT_rich"
complement(17642..17687)
/rpt_family="AT_rich"
17774..17794
/rpt_family="AT_rich"
17926..17970
/rpt_family="AT_rich"
18152..18196
/rpt_family="AT_rich"
complement(19345..21515)
/gene="At2g43920"
/note="synonym: F6B13.5"
complement(join(<19342..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,
21416..21515))
/gene="At2g43920"
complement(join(19586..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,
21416..21515))
/gene="At2g43920"
/codon_start=1
/product="unknown protein"
/protein_id="AAC23401.1"
/db_xref="GI:3212850"
/translation="MAEQNSYSIGNILLPTREBAATFQPVAVAGMDKCMEDGV
TPMDQARATPLIHLHDSALPLGRVLVPGCGHVVMAASPERVVGIDISDKALN
KANETYSPTKAEYFSEFVEDVPTMPSNELFDLIFDYFCAIEPERPAMGSMEL
LKPQGLITLTPMTHEGAPYKVALSGYEDLVVGFRAVSEENPDSIPTRKQGE
KLARKKIN"
23418..25849
/gene="At2g43930"
/note="synonym: F6B13.6; contains a protein kinase domain
profile (PDOC00100)"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
/codon_start=1
/product="putative protein kinase"
/protein_id="AAC23427.1"
/db_xref="GI:3212876"
/translation="MDHRRRLATIKILIRNSSTLLFPKIVPILRLSGGDSRALPYT
PFQCSRALNIGHNTSRSEPHVILQCKRKSVCPSSEFHLCLFGYIDEDERKHF
GHKGVVAREEROLLKNSHPNIVYIDGTVRSDDLNTMMQVPGGSLSLAKXGS
FPEPKHNRVLDNAISHTSYKNIHQGSFKLEIDPRKKGS"
complement(27269..30414)
/gene="At2g43940"
/note="synonym: F6B13.7"
complement(join(<27269..27361,27641..27771,27853..27984,
28336..28397,28479..28537,28635..28784,28865..28981,
29149..29343,30043..>30414))
/gene="At2g43940"
repeat_region
complement(15108..15141)
/rpt_family="AT_rich"
complement(16570..16608)
/rpt_family="AT_rich"
complement(17642..17687)
/rpt_family="AT_rich"
17774..17794
/rpt_family="AT_rich"
17926..17970
/rpt_family="AT_rich"
18152..18196
/rpt_family="AT_rich"
complement(19345..21515)
/gene="At2g43920"
/note="synonym: F6B13.5"
complement(join(<19342..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,
21416..21515))
/gene="At2g43920"
complement(join(19586..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,
21416..21515))
/gene="At2g43920"
/codon_start=1
/product="unknown protein"
/protein_id="AAC23401.1"
/db_xref="GI:3212850"
/translation="MAEQNSYSIGNILLPTREBAATFQPVAVAGMDKCMEDGV
TPMDQARATPLIHLHDSALPLGRVLVPGCGHVVMAASPERVVGIDISDKALN
KANETYSPTKAEYFSEFVEDVPTMPSNELFDLIFDYFCAIEPERPAMGSMEL
LKPQGLITLTPMTHEGAPYKVALSGYEDLVVGFRAVSEENPDSIPTRKQGE
KLARKKIN"
23418..25849
/gene="At2g43930"
/note="synonym: F6B13.6; contains a protein kinase domain
profile (PDOC00100)"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
/codon_start=1
/product="putative protein kinase"
/protein_id="AAC23427.1"
/db_xref="GI:3212876"
/translation="MDHRRRLATIKILIRNSSTLLFPKIVPILRLSGGDSRALPYT
PFQCSRALNIGHNTSRSEPHVILQCKRKSVCPSSEFHLCLFGYIDEDERKHF
GHKGVVAREEROLLKNSHPNIVYIDGTVRSDDLNTMMQVPGGSLSLAKXGS
FPEPKHNRVLDNAISHTSYKNIHQGSFKLEIDPRKKGS"
complement(27269..30414)
/gene="At2g43940"
/note="synonym: F6B13.7"
complement(join(<27269..27361,27641..27771,27853..27984,
28336..28397,28479..28537,28635..28784,28865..28981,
29149..29343,30043..>30414))
/gene="At2g43940"

Alignment Scores:

Pred. No.:	Length:	Matches:
5,23e-22	109741	72
327.00	Matches:	72
82.2%	Conservative:	11
Percent Similarity:	Mismatch:	12
Best Local Similarity:	Indels:	6
Query Match:	Gaps:	3
DB:		15

[illegible]

JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:s.kikuchi@ias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.

COMMENT URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Onedaa, S., Yahagi, W., Suzuki, K., Li, C., Ohnuki, K., Shishiki, T. and
Yamamoto, M.
RAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurisaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Nakikawa, R., Nikkura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tanoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hangaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiroyaka, T., Horii, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Ittoh, M., Kagawa, T., Kanagawa, S., Katoh, H., Kawasaki, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Koude, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakanur, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Osato, N.,
Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazune, N., Sano, H.,
Sasael, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Khalil, S., Tanaka, Y., Tomaru, A., Toyra, T., Wakli, K.,
Yasuishi, A. and Hayashizaki, Y.
Location/Organisms

FEATURES
source
1..1137
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="001-021-P08"

ORIGIN

Alignment Scores:
Pred. No.: 1,51e-13 Length: 1137
Score: 224.50 Matches: 51
Percent Similarity: 65.7% Conservative: 16
Best Local Similarity: 50.0% Mismatches: 24
Query Match: 42.1% Indels: 11
DB: Gaps: 3

US-10-715-129-2 (1-106) x AK060546 (1-1137)

Oy 3 ValG|AArgAanbaAnxLySaAmbeSerPheArgsSerProGluNySeerLygLn 22
Db 688 GTAGGGCTTAGGACGAAGAAGAAAC--CCTTGATACCAGAACCAACGGTTCAAAGAG 744
Oy 23 GluLeuAArgSerPhe-----SerAlaGlnLySarMetMetIleProalaAsn 39
Db 745 CATCGCACACATCAGCATTCAGCAAAATGCTCCGAGAAAGAACCGCGGACATGCAGAT 804
Oy 40 TyrPheSerLeuGlusertPheuleuValGlyLeuthralaSerLeuLeuLleu 59
Db 805 TACTTCAGTATGGAAGGCCGTCTCTCGTCTGCTTCTCCACCAATGATTCATATCACT 864
Oy 60 ProLeuValLeuProProLeuProProProPhetheuLeuLeuValProLleGly 79
Db 865 CCATTGTCCTTCCCCCATGTGCTCCGCGGACATGCTGCTGCTGCCAGTCTGC 924
Oy 80 IleMetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAanaLaenThr 99
Db 925 CTCGCTATCTCTGCTGTGTGTGTGCTGCCTTATGCCA-----ACG 963
Oy 100 AspVal 101

Db 964 GATGTG 969

RESULT 12
AP008217_078/c
WPCOMMENT

Sequence split into 284 fragments LOCUS AP008217 Accession AP008217

Fragment Name	Begin	End
AP008217_000	1	110000
AP008217_001	100001	210000
AP008217_002	200001	310000
AP008217_003	300001	410000
AP008217_004	400001	510000
AP008217_005	500001	610000
AP008217_006	600001	710000
AP008217_007	700001	810000
AP008217_008	800001	910000
AP008217_009	900001	1010000
AP008217_010	1000001	1110000
AP008217_011	1100001	1210000
AP008217_012	1200001	1310000
AP008217_013	1300001	1410000
AP008217_014	1400001	1510000
AP008217_015	1500001	1610000
AP008217_016	1600001	1710000
AP008217_017	1700001	1810000
AP008217_018	1800001	1910000
AP008217_019	1900001	2010000
AP008217_020	2000001	2110000
AP008217_021	2100001	2210000
AP008217_022	2200001	2310000
AP008217_023	2300001	2410000
AP008217_024	2400001	2510000
AP008217_025	2500001	2610000
AP008217_026	2600001	2710000
AP008217_027	2700001	2810000
AP008217_028	2800001	2910000
AP008217_029	2900001	3010000
AP008217_030	3000001	3110000
AP008217_031	3100001	3210000
AP008217_032	3200001	3310000
AP008217_033	3300001	3410000
AP008217_034	3400001	3510000
AP008217_035	3500001	3610000
AP008217_036	3600001	3710000
AP008217_037	3700001	3810000
AP008217_038	3800001	3910000
AP008217_039	3900001	4010000
AP008217_040	4000001	4110000
AP008217_041	4100001	4210000
AP008217_042	4200001	4310000
AP008217_043	4300001	4410000
AP008217_044	4400001	4510000
AP008217_045	4500001	4610000
AP008217_046	4600001	4710000
AP008217_047	4700001	4810000
AP008217_048	4800001	4910000
AP008217_049	4900001	5010000
AP008217_050	5000001	5110000
AP008217_051	5100001	5210000
AP008217_052	5200001	5310000
AP008217_053	5300001	5410000
AP008217_054	5400001	5510000
AP008217_055	5500001	5610000
AP008217_056	5600001	5710000
AP008217_057	5700001	5810000
AP008217_058	5800001	5910000
AP008217_059	5900001	6010000
AP008217_060	6000001	6110000
AP008217_061	6100001	6210000
AP008217_062	6200001	6310000
AP008217_063	6300001	6410000
AP008217_064	6400001	6510000

AP008217_065	6500001	6610000
AP008217_066	6600001	6710000
AP008217_067	6700001	6810000
AP008217_068	6800001	6910000
AP008217_069	6900001	7010000
AP008217_070	7000001	7110000
AP008217_071	7100001	7210000
AP008217_072	7200001	7310000
AP008217_073	7300001	7410000
AP008217_074	7400001	7510000
AP008217_075	7500001	7610000
AP008217_076	7600001	7710000
AP008217_077	7700001	7810000
AP008217_078	7800001	7910000
AP008217_079	7900001	8010000
AP008217_080	8000001	8110000
AP008217_081	8100001	8210000
AP008217_082	8200001	8310000
AP008217_083	8300001	8410000
AP008217_084	8400001	8510000
AP008217_085	8500001	8610000
AP008217_086	8600001	8710000
AP008217_087	8700001	8810000
AP008217_088	8800001	8910000
AP008217_089	8900001	9010000
AP008217_090	9000001	9110000
AP008217_091	9100001	9210000
AP008217_092	9200001	9310000
AP008217_093	9300001	9410000
AP008217_094	9400001	9510000
AP008217_095	9500001	9610000
AP008217_096	9600001	9710000
AP008217_097	9700001	9810000
AP008217_098	9800001	9910000
AP008217_099	9900001	10010000
AP008217_100	10000001	10110000
AP008217_101	10100001	10210000
AP008217_102	10200001	10310000
AP008217_103	10300001	10410000
AP008217_104	10400001	10510000
AP008217_105	10500001	10610000
AP008217_106	10600001	10710000
AP008217_107	10700001	10810000
AP008217_108	10800001	10910000
AP008217_109	10900001	11010000
AP008217_110	11000001	11110000
AP008217_111	11100001	11210000
AP008217_112	11200001	11310000
AP008217_113	11300001	11410000
AP008217_114	11400001	11510000
AP008217_115	11500001	11610000
AP008217_116	11600001	11710000
AP008217_117	11700001	11810000
AP008217_118	11800001	11910000
AP008217_119	11900001	12010000
AP008217_120	12000001	12110000
AP008217_121	12100001	12210000
AP008217_122	12200001	12310000
AP008217_123	12300001	12410000
AP008217_124	12400001	12510000
AP008217_125	12500001	12610000
AP008217_126	12600001	12710000
AP008217_127	12700001	12810000
AP008217_128	12800001	12910000
AP008217_129	12900001	13010000
AP008217_130	13000001	13110000
AP008217_131	13100001	13210000
AP008217_132	13200001	13310000
AP008217_133	13300001	13410000
AP008217_134	13400001	13510000
AP008217_135	13500001	13610000
AP008217_136	13600001	13710000
AP008217_137	13700001	13810000

AP008217_138	13800001	13910000
AP008217_139	13900001	14100000
AP008217_140	14000001	14110000
AP008217_141	14100001	14210000
AP008217_142	14200001	14310000
AP008217_143	14300001	14410000
AP008217_144	14400001	14510000
AP008217_145	14500001	14610000
AP008217_146	14600001	14710000
AP008217_147	14700001	14810000
AP008217_148	14800001	14910000
AP008217_149	14900001	15010000
AP008217_150	15000001	15110000
AP008217_151	15100001	15210000
AP008217_152	15200001	15310000
AP008217_153	15300001	15410000
AP008217_154	15400001	15510000
AP008217_155	15500001	15610000
AP008217_156	15600001	15710000
AP008217_157	15700001	15810000
AP008217_158	15800001	15910000
AP008217_159	15900001	16010000
AP008217_160	16000001	16110000
AP008217_161	16100001	16210000
AP008217_162	16200001	16310000
AP008217_163	16300001	16410000
AP008217_164	16400001	16510000
AP008217_165	16500001	16610000
AP008217_166	16600001	16710000
AP008217_167	16700001	16810000
AP008217_168	16800001	16910000
AP008217_169	16900001	17010000
AP008217_170	17000001	17110000
AP008217_171	17100001	17210000
AP008217_172	17200001	17310000
AP008217_173	17300001	17410000
AP008217_174	17400001	17510000
AP008217_175	17500001	17610000
AP008217_176	17600001	17710000
AP008217_177	17700001	17810000
AP008217_178	17800001	17910000
AP008217_179	17900001	18010000
AP008217_180	18000001	18110000
AP008217_181	18100001	18210000
AP008217_182	18200001	18310000
AP008217_183	18300001	18410000
AP008217_184	18400001	18510000
AP008217_185	18500001	18610000
AP008217_186	18600001	18710000
AP008217_187	18700001	18810000
AP008217_188	18800001	18910000
AP008217_189	18900001	19010000
AP008217_190	19000001	19110000
AP008217_191	19100001	19210000
AP008217_192	19200001	19310000
AP008217_193	19300001	19410000
AP008217_194	19400001	19510000
AP008217_195	19500001	19610000
AP008217_196	19600001	19710000

Alignment Scores:

Pred. No.:

1.73e-11

Length:

110000

Score:

224.50

Matches:

51

Percent Similarity:

65.7%

Conservative:

16

Best Local Similarity:

50.0%

Mismatches:

24

Query Match:

42.1%

Indels:

11

DB:

15

Gaps:

3

US-10-715-129-2 (1-106) x AP008217_078 (1-110000)

Ory	23	GlutelinArygserphe-----SerAlaglnlyAArgMetMetIleProAlaAsn 39 ::: ::: ::: :::
Db	58065	CATCTGCAGCATACGATCATACCAAATGCTGCCGAAGAAGAACCGCCGACATCGAAT 580066
Oy	40	TyrPheSerLeuGluIuSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuIleLeu 59 ::: ::: ::: ::: :::
Db	58005	TACTTCAGTATCAGAGCGCTCTCCGTGGCTGCTTCTTCTCACCATGTCATTGCTCACTT 579466
Oy	60	ProLeuValLeuProProIleuR-Pro-ProporProPheMetLeuLeuValProlleoly 79 ::: ::: ::: ::: :::
Db	57945	CAATTGGTGTTCCCCCATGCTCCGCCGACATCGCTCTGTGTGTGCAGCTTGC 578866
Oy	80	IIMetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAnaAlaSerThr 99 ::: ::: ::: ::: ::: :::
Db	57885	CTGCTATCTCTGCTGTTGTTGCTGCGCTTCATGCCA-----ACG 578476
Oy	100	AspVal 101 ::: ::: ::: ::: :::
Db	57846	GATGTG 57841
RESULT_13		
LOCUS	AC137924	
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 11 clone	
ACCESSION	OSJNBA0095X08, complete sequence.	
VERSION	AC137924	
KEYWORDS	AC137924.3 GI:46849661	
SOURCE	HMG.	
ORGANISM	Oryza sativa (japonica cultivar-group)	
	Oryza sativa (japonica cultivar-group)	
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
	Elephantodae; Oryzace; Oryza.	
REFERENCE	1 (bases 1 to 180999)	
AUTHORS	Buell,C., Yuan,Q., Quyang,S., Liu,J., Wang,A., Maiti,R., Lin,H., Zhu,W., Hamilton,J., Jones,K., Tallon,L., Feldblum,T., Tsaltrin,T., Bera,J., Kim,M., Jin,S., Padrosh,D., Vuong,H., Overton II,L., Reardon,M., Weaver,B., Johri,S., Lewis,M., Ulteckbeck,T., Van Aken,S., Wortman,J., Hase,B., Koo,H., Zisman,V., Haiso,J., Tobes,S., de Vazelles,A., White,O., Salzberg,S. and Frazer,C. Unpublished	
JOURNAL	2 (bases 1 to 180999)	
REFERENCE	Buell,R.	
AUTHORS	Direct Submission	
TITLE	Submitted (04-DEC-2002) The Institute for Genomic Research, 9712	
JOURNAL	Medical Center Dr. Rockville, MD 20850, USA	
REFERENCE	3 (bases 1 to 180999)	
AUTHORS	Buell,R.	
TITLE	Direct Submission	
JOURNAL	Submitted (10-APR-2003) The Institute for Genomic Research, 9712	
REFERENCE	4 (bases 1 to 180999)	
AUTHORS	Buell,R.	
TITLE	Direct Submission	
JOURNAL	Submitted (29-APR-2004) The Institute for Genomic Research, 9712	
REFERENCE	5 (bases 1 to 180999)	
AUTHORS	Buell,R.	
TITLE	Direct Submission	
JOURNAL	Submitted (19-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	6 (bases 1 to 180999)	
AUTHORS	Buell,R.	
TITLE	Direct Submission	
JOURNAL	Submitted (26-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	7 (bases 1 to 180999)	
AUTHORS	Buell,R.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases 1 to 180999)	
AUTHORS	Medical Center Dr. Rockville, MD 20850, USA, rbuell@tigr.org	
JOURNAL	Submitted (27-APR-2005) The Institute for Genomic Research, 9712	
REFERENCE	8 (bases	

AUTHORS Buell, R.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2005) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
COMMENT On Apr 29, 2004 this sequence version replaced gi:29725796.
Address all correspondence to: rice@tigr.org

BAC clone OSJNBa0095K08 is from Oryza sativa chromosome 11. The orientation of the sequence is from Sp6 to T7 end of the BAC clone.
Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (<http://www.softberry.com/>), GenScan and Genscan+ (Chris Burge, <http://CCR-081.mit.edu/GENSCAN.html>), GenemarkHM (Mark Borodovsky, <http://genemark.biology.gatech.edu/Genemark/>), and Geneslicer (Michael Perera and Steven Salzberg, [contact mperera@tigr.org](http://contact.mperera@tigr.org)), searches of the complete sequence against a peptide database and the plant EST database at TIGR (<http://www.tigr.org/cdb/cgi.shtm>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/CRNAScan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

FEATURES
SOURCE

```
1. 180999
  /organism="Oryza sativa (japonica cultivar-group)"
  /mol_type="genomic DNA"
  /cultivar="Nipponbare"
  /db_xref="taxon:39947"
  /chromosome="11"
  /clone="OSJNBa0095K08"
  /note="japonica cultivar-group"
  complement(join(<771..2739,5345..>6132))
  /product="NB-ARC domain, putative"
  /note="go function: molecular function [go:0003674];
  go process: response to biotic stimulus [go:0009607]"
  /codon_start=1
  /product="NB-ARC domain, putative"
  /protein_id="AAJ96748.1"
  /db_xref="gi:62734638"
  /translation="MEAAIYASATGWRSLAKLTNLTGGHYKLLKWLRRMEPLSE
  LRMSITLEREDYOKAHPOKMDRVRRLAYIDECIDPFIQDLSKAKYFGOK
  LLSARRIGHMIRBLKARWESERKRRYMLDGLSGSVKVPRLSALVYDRLV
  GIDAPREIIGRLDKRRSASAKQVMTISVCGGLGKTLTANQICKIKGFCALF
  ASVPQNTKVLNTNLSQVATTAVEDBOAILNKRLBYSDRYIVIVDIDCMQI
  WKPEICALPVKCSRRIITTRIDHIAKCCSHGDYIEMKPGVIDSKLFPKRIE
  DPERRRPNOLTEVEREILKKCGPLAIISSILKSGKMDKMDKYSLSSTLR
  TPDLEMTWLSISYSDLPNHTKCLYISFPFGYEINERLVSMTIASFTYKKG
  QNPEVSDSYFNELVNRSLLIQPANIKPDGQTNACRVDTVDHDFVMSVEENFTLG
  GSKLPVSHGKVRSLISONGIENIYVTSQVRSLLTFAVEMPSLGSFMTLVG
  LLDIEDCYALBDHMLTNLERLVOLRYSIRTSPIELPQIGQLQYLFTKEIGLATVE
  LPSTIGRLKSLVLFVYHVKLPKEISNMALBELTSPSALMSPDFTKEIGQLTNRR
  VARYICDGSFKDAGSCLENTASSLCNIGYVNHSLFVDINGYEDNPSIDTQPP
  SRLARPSIDRKPINKIPNWKGLINBELVLYNKIQWDFELGMPALSLTITS
  NTLAQRIISGFHSITPKFYCNPAQITPDAGSLKLECLDVIMNVNTKSSGSDP
  FGIQYLTNLRNVYIQDONGSTGELBLAAKASIKSVNKLPGQPKMLSTLTENMLVH
  EKAN"
```

```
/protein_id="AAJ96748.1"
/db_xref="gi:62734638"
/translation="MEKLTGVKRGAGSDGSPRADSGRGGSRDGRGSGRDERGVYD
GPFYDPDFRLILAM"
9150..9191
  /note="CT-rich"
  9211..9231
  /note="GC rich"
  9346..9421
  /note="microsatellite"
  /rpt_type=tandem
  /rpt_unit="ccg"
  10046..10131
  /note="microsatellite"
  /rpt_type=tandem
  /rpt_unit="ccg"
  complement(join(<13607..13947,15379..>215640))
  /product="transposable element protein, putative"
  complement(join(13607..13947,15379..15640))
  /note="containing Pfam profile(s): PF00098"
  /codon_start=1
  /product="transposable element protein, putative"
  /protein_id="AAJ96749.1"
  /db_xref="gi:62734640"
  /translation="MDNWILMOSFYNGITPSSODHLIAAGAPFSTYVGAIGLIEK
  WSNITAGANNDRPPIVACTLLISLNDITRNTIPSGARNIHDTIKETRLAKTDL
  LMKRLDQDSQPGQTVKALDSHISCVGCMGSGNDCLTREBAMFWMGNNGYR
  PHKGGQNNQDPFPFHGGNNNGNPNQPSLRHTVPPDRR"
  17831..18204
  /note="microsatellite"
  /rpt_type=tandem
  /rpt_unit="ta"
  complement(join(<18256..18438,18518..18616,18719..18791,
  18879..18989,19086..19259,19371..19481,19671..20119,
  20215..>21636))
  /product="transposon protein, putative, CACTA, En/Spm
  sub-classes"
  complement(join(18256..18438,18518..18616,18719..18791,
  18879..18989,19086..19259,19371..19481,19671..20119,
  20215..21636))
  /note="similar to TIGR Oryza Repeat Database entry
  ORSGTERRN00200008"
  /codon_start=1
  /product="transposon protein, putative, CACTA, En/Spm
  sub-classes"
  /protein_id="AAJ96750.1"
  /db_xref="gi:62734641"
  /translation="MADRNEQLIYDTIAEGSSQYVMEEGENDPNQYLINEEGVBD
  ABGNOQGHVRDVBGNDBEASGSQSPVCGQRARGOGAKLEGRHITTEVEBGRP
  SAPLEAKNTVNRHNGVNRDNVPSTVYWRTRARDEHSFVPSSEKEMLTMLFTF
  TTPVGTDKVRYWTLKKMAEQPOSFKDLPQKYLKQTPNFDPLRLDMDHMFVAY
  KTBEGOCAMERKNENAKKKYHHLSSGGYSVAMPKMEIASILRGLPEAPANRP
  ERSKFYVANGTLLNPDGSLVGEDQIRBAARLIDVVBASQGTPTPDGRDCLTLL
  LQNDHEHGRTRGGVTPWKIGKEDIDTYSRMSRSDTAKTYDLEFRRSYSLSD
  EYARKYDEMAHRSHPDPTTPPAVVSQGNHSCASGTQVQSOMDAMQTDDEST
  CPVDDITDYPCEBLIIPKULSIKVASGMAIPDPSCTYICRPIPAYSKVEVLVIG
  AYBDELRDYGGEETRIIDTSLIILMRGTSLSLGDKRLVHLLRLHLRLIILH
  RLMLRORHLRLRORHLRLIOLRLIGLILHLRLPRLHLOSQOPOLHLPTOG
  RPSSCOGDILMNFKGIYELIYOLDALDVIMSCWIMELORARORGVDPDGLDPR
  VNVAMLDQVQOEREDNIVHLKQHYTPIILPNTREPHYVLLFDLSACTVNYVDSM
  DKRSTFDKVFBLIDRAWITFRHLVRNMRRLRRKRPCKAKQKQSGTINCGYTWCK
  CHCIADQITTRKLDPTIRMNNTLTHKEPIAAVQOLMGFINBIIDPKGEFYVDGNT
  IHSLSLELATYTTTSSK"
  19723..19882
  /note="GA-rich"
  21444..21528
  /note="CT-rich"
  22688..23029
  /note="microsatellite"
  /rpt_type=tandem
  /rpt_unit="ctcg"
  23067..23203
  /note="microsatellite"
```

```

repeat_region /rpl_type=tandem
                /rpl_unit="cgs"
                /note="microsatellite"

Alignment Scores:
Pred. No.: 2.89e-11 Length: 180999
Score: 224.50 Matches: 51
Percent Similarity: 65.7% Conservative: 16
Best Local Similarity: 50.0% Mismatches: 24
Query Match: 42.1% Indels: 11
DB: 15 Gaps: 3

US-10-715-129-2 (1-106) x AC137924 (1-180999)

Qy 3 ValGlyArgAsnAnaGlyArgAsnMetSerPheArgSerProGlyIlySerIlyGln 22
    |||||
Db 139826 GTAGGTCTCTAGGACAGAGAGAGAAC--CTCTGTATACCAGACCAAGCGTTCAAGAGAG 139882
    |||||

Qy 23 GluLeuArgArgSerPhe-----SerIaGlnIlyArgMetMetIleProAlaAsn 39
    |||||
Db 139883 CATCTGAGCATGACATGACATGACCAATGCTGCCAGAGAAAGACCCGCCGACATCAAT 139942
    |||||

Qy 40 TyrPheSerLeuGluSerLeuPheLeuValGlyLeuThraIaSerLeuLeu 59
    |||||
Db 139943 TACTTCAGTATCGAGGCGTCTCTGCTGCTCTCTCCATGCATGCTATGCTATCTT 140002
    |||||

Qy 60 ProLeuValLeuProProLeuProProProPheMetLeuLeuValProIleGly 79
    |||||
Db 140003 CCATGAGTGTCTCCCGCATGTCCTCCGCCGATGCTGCTGCTGCTGCCAGTGTGC 140062
    |||||

Qy 80 IleMetValLeuLeuValValIleValAlaPheMetProSerSerHisSerAsnAlaAsnThr 99
    |||||
Db 140063 CTGCTCATCTCGCTGCTGCTGCTGCTGCTGCTATGCA-----ACG 140101
    |||||

Qy 100 AspVal 101
    |||||
Db 140102 GATGTG 140107
    |||||

RESULT 14
AKI02641 1302 bp mRNA linear PLN 24-JUL-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J033100F02, full
DEFINITION Insert sequence.
ACCESSION AKI02641.1 GI:32987850
VERSION FLI_CDNA; CAP trapper.
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Bkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ORGANISM Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzeae; Oryza.

REFERENCE
1 The Rice Full-length cDNA Consortium, National Institute of
  Agricultural Sciences Rice Full-length cDNA Project Team,
  Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
  Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotte,I.,
  Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
  Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
  Science Genome Sequencing & Analysis Group, Ohtomo,Y., Murakami,K.,
  Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
  Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
  Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J.,
  Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Mura,J.,
  Kusunegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
  Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
  Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
  Kagawa,I., Kondo,S., Kono,H., Miyazaki,A., Otsu,N., Ota,Y.,
  Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
  Yoshino,M. and Hayashizaki,Y.
  Collection, mapping, and annotation of over 28,000 cDNA clones from
  japonica rice
  Science 301 (5631), 376-379 (2003)

```

```

PUBMED 12689764
REFERENCE 2 (bases 1 to 1302)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
  Fujimura,T., Fukuda,S., Hangaki,T., Hara,A., Hashidume,W.,
  Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hirooka,T.,
  Hori,F., Hotte,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
  Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
  Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
  Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
  Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Kono,H., Koude,M.,
  Koya,S., Kurihara,C., Kurosaki,T., Kusunegi,T., Li,C., Lu,M.,
  Masuda,H., Matsubara,K., Matsuyama,T., Mura,J., Miyazaki,A.,
  Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
  Namiki,T., Narikawa,R., Niihara,J., Niihara,J., Nomura,K.,
  Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
  Otsu,N., Ota,Y., Ohtomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
  Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
  Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
  Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
  Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
  Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
  Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
  Yoshimura,A.
  Direct Submission
  Submitted (27-AUG-2002) Shoeni Kikuchi, National Institute of
  Agricultural Sciences, Department of Molecular Genetics, Head of
  Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki
  305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
  Tel:81-29-858-7007, Fax:81-29-858-7007)
  This clone is one of the 28k full-length cDNA clones from japonica
  rice.
  URL: http://cdna01.dna.affrc.go.jp/cDNA/
  NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
  Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
  Ishikawa,M., Yamada,H., Ooka,H., Hotte,I., Kojima,K., Namiki,T.,
  Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
  Yamamoto,M.
  FAIS Genome Sequencing & Analysis Group: Ohtomo,Y., Iida,Y.,
  Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
  Kodama,T., Kurosaki,T., Kusunegi,T., Lu,M., Masuda,H., Mura,J.,
  Mizuno,K., Narikawa,R., Niihara,J., Oka,M., Ryu,R., Sugano,S.,
  Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
  Yoshimura,A., Matsubara,K. and Murakami,K.
  Genome Exploration Research Group in Riken Genomic Sciences Center
  and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
  Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hangaki,T.,
  Hara,A., Hashidume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
  Hirooka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
  Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
  Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Kono,H., Koude,M.,
  Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
  Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohtomo,Y., Otsu,N.,
  Oka,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
  Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
  Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
  Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
  Yasunishi,A. and Hayashizaki,Y.
  Location/Qualifiers
  1. 1302
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="mRNA"
    /cultiar="Nipponbare"
    /db_xref="taxon:39947"
    /clone="J033100F02"

FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 6.78e-12 Length: 1302
Score: 209.00 Matches: 40
Percent Similarity: 90.0% Conservative: 14
Best Local Similarity: 66.7% Mismatches: 6
Query Match: 39.2% Indels: 0
DB: 15 Gaps: 0

```

Y	40	TyrPheSerLeuGluSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuIleLeu	59	
Db	872	TACTTCTCGAGGTGCCTCTCCTCTGTCGTGGTCGACGAGCGCCCTCGCTGATCTCTC	931	
Oy	60	ProteuValLeuProPheLeuProPheProPhePheMetLeuLeuValProIleGly	79	
Db	932	CCGCTGTCGTCCGCCCGCTCCCGCGCGCTGATGCTGATGCTGATGCTGATGCTGCG	991	
Oy	80	IleMetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAnthr	99	
Db	992	ATGCTGTCGTCTCTCTGCGCTGCTTTATTATCCGACGACGACGCTCTCTCTCTCTC	1051	
RESULT 15				
LOCUS	OSUNB00249	46422 bp	DNA linear	
DEFINITION	Oryza sativa genomic DNA, chromosome 4, BAC clone: OSUNB0058G03,			
ACCESSION	AL731606			
VERSION	AL731606.2	GI:32482924		
KEYWORDS	HTG.			
SOURCE ORGANISM	Oryza sativa (japonica cultivar-group)			
REFERENCE AUTHORS	Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehharitoidae; Oryzaceae; Oryza. 1 Feng,Q., Zhang,Y., Hao,P., Mang,S., Pu,G., Huang,Y., Li,Y., Zhu,J., Liu,Y., Hu,X., Jia,P., Zhang,Y., Shao,Q., Ying,K., Yu,S., Tang,Y., Weng,Q., Zhang,L., Lu,Y., Mu,J., Liu,Y., Zhang,J.S., Yu,Z., Fan,D., Lin,X., Lu,T., Li,C., Wu,Y., Sun,T., Lei,H., Li,T., Hu,R., Guan,J., Wu,M., Zhang,R., Zhou,B., Chen,Z., Chen,L., Jin,Z., Wang,R., Yin,H., Cai,Z., Ren,S., Lv,G., Gu,W., Zhu,G., Tu,Y., Jia,J., Zhang,X., Chen,Z., Kang,H., Chen,X., Shao,C., Sun,Y., Hu,Q., Zhong,X., Zhang,W., Wang,L., Ding,C., Sheng,H., Gu,J., Chen,S., Ni,L., Zhu,F., Chen,W., Lan,L., Lai,Y., Cheng,Z., Gu,M., Jiang,J., Li,J., Hong,G., Xue,Y. and Han,B. Sequence and analysis of rice chromosome 4 Nature 420 (6913), 316-320 (2002) 12447439			
TITLE	JOURNAL	PUBMED		
REFERENCE	AUTHORS			
	2	Han,B., Feng,Q., Huang,Y.C., Li,Y., Zhu,J.J., Zhao,Q., Hu,X., Liu,Y.L., Yu,Z., Chen,L., Fan,D.L., Weng,Q.J., Zhang,L., Lu,Y.Q., Mu,S.L., Liu,X.H., Lu,T.T., Zhang,Y.J., Li,C., Li,T., Zhang,Y., Hu,H., Jia,P.X., Qian,Y.M., Ying,K., Zhou,B., Chen,Z.H., Hao,P., Zhang,L., Wu,M., Zhang,R.Q., Guan,J.P., Pu,G., Wang,S.Y., Ren,S., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F., Tu,Y.F., Jia,X.Y., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H., Chen,X.Y., Shao,C.Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W., Wang,L.J., Ding,C.W., Sheng,H.H., Gu,J.L., Chen,S.T., Ni,L., Zhu,F.H. and Hong,G.F. Direct Submission Submitted (04-MAY-2002) Han Bin, National Center for Gene Research, Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233, CHINA. E-mail enquiries: bhanncgr.ac.cn. Clone requests: bhanncgr.ac.cn Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: OSUNB0058G03. On Jul 9, 2003 this sequence version replaced gi:20451715. Web site: http://www.ncgr.ac.cn ----- Summary Statistics ----- Assembly program: phrap		
REMARK				
COMMENT				

This is a complete sequence.

Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (<http://www.softberry.com/>), genscan (<http://CCR-081.mt.edu/genScan.html>), GenemarkHM (<http://genemark.biology.gatech.edu/genemark/>), tRNAscan-SR (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscanner-SE/>), searches of the complete sequence against NCBI nonredundant protein database (<ftp://ncbi.nlm.nih.gov/blast/db>) and the EST database at NCRG.

FEATURES

source

Location/Qualifiers

1..46422

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="japonica"

/cultivar="Nipponbare"

/sub_species="japonica"

/db_xref="taxon:39947"

/chromosome="4"

/clone="OSUNBA0058G03"

/clone_1b="CUG1-OSUNBA"

complement (1144..2100)

/gene="OSUNBA0058G03.1"

complement (1144..2100)

/gene="OSUNBA0058G03.1"

/codon_start=1

/protein_id="CAE02427.2"

/db_xref="GI:38344583"

/db_xref="UniProt/TREMBL:O7F9D9"

/translation="MRHPTPTNIGVTVTVLSIVSLGLCLCHSLNPOLIKRR
RNYIQADBOLESTFNGFWLTRLITLILVALMWGEVLRITFVNGEGRISDQWQAV
CKRYISNLGASRGLEFLALFPLSALOQEVGALKRNQRTI CAFWLCPSLIL
EACVIFPGHIASNDGQTSKATXWYSSASVHDGACVCPILSIFELGTFVTLTY
VIFGQGITSLVINKGRLRRIRIYMLIFATGILLPRMELGSVLPWPEIIVHSLVFS
FLVIMIAMGIYLVIVFPALTEFVRNQHILBLQSHSLAL"

complement (6127..6681)

/gene="OSUNBA0058G03.2"

complement (6127..6681)

/gene="OSUNBA0058G03.2"

/codon_start=1

/protein_id="CAD41044.1"

/db_xref="GI:21742878"

/db_xref="GOA:Q7XUS1"

/db_xref="UniProt/TREMBL:IPR001471"

/db_xref="InterPro:IPR001471"

/translation="MASAAAAAGCGCGRGVRRKMGKWSBIRPOETKTIRMLGSP
SABMAAHDVAALRLRGDQALNPSRVRLRPASSRGDIRAAAEADYRREP
ALVAVRGAAAGELAAVRWAGLEVEEQLGSDDEFVDSFLMAEMABMLDPP
VMAVDSVEMEGPHCMAHGLMDAC"

17958..18485

/gene="OSUNBA0058G03.3"

17958..18485

/gene="OSUNBA0058G03.3"

/codon_start=1

/protein_id="CAE02428.1"

/db_xref="GI:32482926"

/db_xref="GOA:Q7XRMI"

/db_xref="InterPro:IPR007087"

/db_xref="UniProt/TREMBL:Q7XRMI"

/translation="MEFVWEADAAGAGSEGLTSLSLQSPPRFQALFSCCYCPKRF
RSSALGCHONAHKLCNRLARRGEAAASIAAAAAAASGDQOQGRTTAAAAA
AVLAGSAPAPAAARAAADLGAAGWGAGRGRRPAHHRHLMQGVSSGSAAGRGGR
NGRLADEMDLSKTL"

24188..29072

/gene="OSUNBA0058G03.4"

join(24188..24309,25603..25678,25778..25848,25983..264
26560..26720,26810..27013,27773..28007,28477..29072)

/gene="OSUNBA0058G03.4"

/codon_start=1

/protein_id="CAE02429.1"

/db_xref="GI:32482927"

/db_xref="UniProt/TREMBL:Q7XRNO"

/translation="MQRGRLHLSRGSKRLADAGGGGDDDRAPKRPVVALASVIV
EALKVDSLOKLSLEPLIRVSVSEVRLALAKGPAATAPARIGRSPPKRIEPPGCI
NLOPFSRSLSLPLTGGKVEGGGAAIHVLLDANGRVAVTSGPSFALDVLVEGG
DRNKEQDEBMTBEFEENHIVKERBGRPLLTGDLQVTLKSGVCTIGLITTDNSWR
SRFRILRLASVSGFCVEHVRKEATFAFTVKDRHGLYKQYIPALKDQDWRLEKIC
DGAFHKLNSNGIYTVBFPLQLVLRDQKRLTIGSNMSNMWESLVEAKTCVLSGK
HIYIYSDARSVGAIPNNIYFETGLADDOYISAEINSENQRLFADTLVQAAYDMIN
VVEYDGLSLRFKOKKSVTTRSDTAASVYSSYSGTSHKQLTGCPVNIQSSMS
SMASGTRMNSNGOARVYANPDIQSITMVDMSLARPEQFAGSITQOTASRS
NMLALGTQOONFBSALGSGMOPSLNPEDDMSRLQBNRGVDYIMEIRVASHSI
LENEEDQOQMLRILSMGSSANNMHGGSFPEMSPAPATNYEDDRPBGKAVVGL

GenCore version 5.1.7
Copyright (c) 1993 - 2006 BioCeleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 30, 2006, 06:39:18 ; Search time 493 Seconds
(without alignments)
1432.977 Million cell updates/sec

Title: US-10-715-129-2

Perfect score: 533
Sequence: 1 MDVGRNRRKMKSRSSPEKS.....LAFWPSHSNANDVTGTFM 106

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10715129/rnat.27012006.154744.14557/app.query.fasta.1.263
-DB=N_Geneseq -OFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-LOCAL=GN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HBASEIZE=500 -MINLEN=2000000000
-USRR=US10715129.OCGN.1.1.1096.0/rnat.27012006.154744.14557 -NCPU=6 -ICPU=3
-NO_MMP -NEG_SCORES=0 -WAIT -DSPLOCK=100 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-THRAUS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21.*

1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	533	100.0	732	14	ADZ67285 Arabidops
2	327	61.4	336	6	ABZ13462 Arabidops
3	327	61.4	336	8	ADA67786 Arabidops
4	327	61.4	1005	3	AAC53386 Arabidops

5	235	44.1	650	3	AAC51918	AAC51918 Arabidops
6	224.5	42.1	1137	14	ABE67601	Reb67601 Rice geno
7	215	40.3	298	11	ACL13088	ACL13088 Rice abio
8	197.5	37.1	978	11	ACL29210	ACL29210 Rice abio
9	195	36.6	665	13	ADK49328	ADK49328 Plant ful
10	193	36.2	990	13	ADK64606	ADK64606 Plant ful
11	193	36.2	1157	13	ADY10467	ADY10467 Plant ful
12	181	34.0	1055	13	ADT18995	ADT18995 Plant ful
13	165.5	31.1	939	13	ADT17295	ADT17295 Plant ful
14	154.5	29.0	936	13	ADK37027	ADK37027 Plant ful
15	149.5	28.0	366	13	ADK64525	ADK64525 Cotton cd
16	144	27.0	577	3	AAC38794	AAC38794 Arabidops
17	144	27.0	2000	6	ABZ16334	ABZ16334 Arabidops
18	144	27.0	2000	8	ADA68185	ADA68185 Arabidops
19	140	26.3	473	8	ACC60011	ACC60011 Rice leaf
20	132	24.8	647	13	ADT17651	ADT17651 Plant cdn
21	121.5	22.8	501	3	AAC54264	AAC54264 Arabidops
22	115	21.6	529	3	AAC53305	AAC53305 Arabidops
23	106	19.9	1863	5	AAS80665	AAS80665 DNA encod
24	105	19.7	2010	13	ADU01858	ADU01858 Novel hum
25	104	19.5	202251	11	ACN44504	ACN44504 Mouse gen
26	99.5	18.7	110000	8	ABX16390_1	ABX16390_1 Continuation (2 of
27	99	18.6	43229	11	ACN44192	ACN44192 Mouse gen
28	99	18.6	235962	14	ADZ12926	ADZ12926 Murine ca
29	98.5	18.5	1197	5	AAS89331	AAS89331 DNA encod
30	98.5	18.5	4128	5	AAS88553	AAS88553 DNA encod
31	97	18.2	110000	8	ABX16390_3	ABX16390_3 Continuation (4 of
32	96.5	18.1	575	4	AAI17342	AAI17342 Probe #72
33	96.5	18.1	575	4	ABA62258	ABA62258 Human fce
34	96.5	18.1	575	4	AAI42236	AAI42236 Probe #10
35	96.5	18.1	575	4	ABZ12604	ABZ12604 Probe #80
36	96.5	18.1	575	4	AAK36486	AAK36486 Human bon
37	96.5	18.1	575	4	AAK10591	AAK10591 Human bra
38	96.5	18.1	575	4	ABS36136	ABS36136 Human liv
39	96.5	18.1	575	6	ABS10483	ABS10483 Human gen
40	96.5	18.1	30306	11	ACN44196	ACN44196 Mouse gen
41	96	18.0	287	4	AAI26561	AAI26561 Probe #16
42	96	18.0	287	4	ABA74794	ABA74794 Human fce
43	96	18.0	287	4	AAI55319	AAI55319 Probe #24
44	96	18.0	287	4	ABA39495	ABA39495 Probe #17
45	96	18.0	287	4	AAK49450	AAK49450 Human bon

ALIGNMENTS

RESULT 1	
ADZ67285	
ID	ADZ67285 standard; cDNA; 732 BP.
XX	
AC	ADZ67285;
XX	
DT	28-JUL-2005 (first entry)
XX	
DE	Arabidopsis thaliana auxin-inducible gene ARGOS.
XX	
KW	cloning; plant breeding; auxin; development; transformation; gene; ss.
XX	
OS	Arabidopsis thaliana.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	388..708
FT	/*tag= a
FT	/product= "ARGOS protein"
XX	
PN	US2005108793-A1.
XX	
PD	19-MAY-2005.
XX	
PP	18-NOV-2003; 2003US-00715129.
XX	
PR	18-NOV-2003; 2003US-00715129.
XX	
PA	(TEMA-) TEMASEK LIFE SCI LAB.

XX ADA67786;
AC
XX 20-NOV-2003 (first entry)
DT
XX Arabidopsis thaliana gene, SEQ ID 23.
DE
XX Plant; bacterial infection; fungal infection; viral infection; ds.
KM
XX Arabidopsis thaliana.
OS
XX
XX WO2003000898-A1.
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PA
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175296/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
PS
XX Claim 6; SEQ ID NO 23; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 336 BP; 72 A; 74 C; 68 G; 122 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1.26e-28 Length: 336
Score: 327.00 Matches: 72
Percent Similarity: 82.3% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: Gaps: 3

US-10-715-129-2 (1-106) x ADA67786 (1-336)
QY 1 MetAspVal---GlyArgAsnAsnArgLysAsnMetSerPheArgSerSer----- 16
Db 1 ATGAGCGTGAAGAGATCATTAACCGGAAAAACAGATTTTCGTGCTTCAGCTCAGCT 60
QY 17 ProGluLysSerLysGlnGlnLeuArgArgSerPheSerIaGlnLys-----ArgMet 34
Db 61 CCGATTATGGGGAAGCAGAAATTGTTTCGACATTGCTGTCGAGAACAGTCCAAAGAGG 120
QY 35 MetIleProIaAsnArgPheSerLeuGlnSerLeuPheLeuLeuValGlyLeuThrIa 54
Db 121 CTATATATAGCGAGATTAATCTTCAGTTTGAATCAATGATGTTGCTTGTCTCAGCA 180
QY 55 SerLeuLeuIleLeuProLeuValLeuProLeuProProProProPheMetLeuLeu 74
Db 181 TCTCTCTGATCTTAACCGTGAATCTTCCACCATTCCTCCCTTTATATGCTGCTT 240
QY 75 LeuValProIleGlyIleMetValLeuLeuValValLeuAlaPheMetProSerSerHis 94

Db 241 TTGATTCCTATTCGATTAACGATTTTCCTATATGATTCCTTCATACCTTCCTATAT 300
QY 95 Ser 95
Db 301 TCC 303

RESULT 4
AAC53386
ID AAC53386 standard; DNA, 1005 BP.
AC
XX AAC53386;
AC
XX 18-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 74346.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-01377502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145182P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147182P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151390P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158022P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-715-129-2 (1-106) x AAC53386 (1-1005)

Length:

Matches:

Conservative:

Mismatch:

Indels:

Gaps:

1005

72

11

12

6

3

QY 1 MetAspVal---(G)ArgAsnAsnArgIlyAsnMetSerPheArgSerSer----- 16
Db 617 ATGACCGAGAGAGATCATACCGGAAACAGAGTTTCTGGTTCACCTCAGCT 676
QY 17 ProGluIlySerIlySGInGluLeuArgSerPheSerIaGlnIlyS-----ArgMet 34
Db 677 CCGATTATGGGAGGACCAAGATTGTTTCGACATTTGCTGCGACAGACAGTCCAAGAGG 736
QY 35 MetIleProAlaAsnIlyrPheSerIeUGInSerIeupheIeuleuValGlyIleuThraIa 54
Db 737 CTATATCAGCGAGATTACTTCAATTAGATCAATGATGTTGCTGTTGCTCAGCA 796
QY 55 SerIeuleuIleuProIeuleuValIeuproIeuproIeuproIeupheMetIeuleu 74
Db 797 TCTCTGTGATCTTACCGTTGATCTTCACCATTTGCTCTCTCTTATATGCTGCTT 856
QY 75 LeuValProIleGlyIleMetValIeuleuValIleuAlaPheMetProSerSerHis 94
Db 857 TTGATTCTTATGGGATTATGCTTTTCTTATGCTTCTTCTTCAATCCTTTCTTAT 916
QY 95 Ser 95
Db 917 TCC 919
RESULT 5
ID AAC51918 standard; DNA; 650 BP.
XX AAC51918;
AC AAC51918;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 70014.
DE Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-00301439.
PF 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131448P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135333P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139482P.
PR 18-JUN-1999; 99US-0139484P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140689P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145221P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.

```

PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149175P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 24-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158359P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.

```

```

PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1,35e-17 Length: 650
Score: 235.00 Matches: 53
Percent Similarity: 71.7% Conservative: 13
Best Local Similarity: 57.6% Mismatches: 14
Query Match: 44.1% Indels: 12
DB: 3 Gaps: 3

US-10-715-129-2 (1-106) x AAC51918 (1-650)

Qy 13 PheArgSerSerProGluYsSerIySGInGluLeuArg-----ArgSerPheSerAla 30
Db 318 TATCTATCGAGTCTGTAAGAAACAGAGATCAACATTTCAGCTTGAGGAGATGTTGGCA 377
Qy 31 GlnIlyArgMet-----MetIleProAlaAntyPheSerIeu 43
Db 378 GATCAACTGTAGAAACATGCTATTATTCAGGTTGATCTTGT-----TTTATA 428
Qy 44 GluSerIeuPheIeuLeuValGlyIeuThrAlaSerIeuLeuValIeuProIleuValIeu 63
Db 429 GACTCAATGGTTGGTGTGGTTCACAGATCTCTTGAATCTTGAATCCGTTGATTTCTT 488
Qy 64 ProPheIeuProProProPheMetIeuLeuValProIleGlyIleMetValIeu 83
Db 489 CACACATTCGCTCTCTCTCTCTTTATAGCTCTTTGAATCTTGAATCCGTTGATTTG 548
Qy 84 IeuValValIeuAlaPheMetProSerSerIleSer 95
Db 549 CTTATGTTCTTCTTCTTTCATGCTTCTTGAATTC 584

RESULT 6
AEB67601
ID AEB67601 standard; DNA; 1137 BP.
XX
AC AEB67601;
XX
DT 22-SEP-2005 (first entry)
XX
DE Rice genome derived DNA sequence, SEQ ID 2746.
XX
KW transcription; gene regulation; transgenic plant; RNA interference;
XX transformation; antibody; ds.
XX
OS Oryza sp.
XX
PN JP2005185101-A.
XX
PD 14-JUL-2005.
XX
PF 11-DEC-2002; 2002JP-00383870.
XX
PR 30-MAY-2002; 2002JP-00203269.
XX
PA (DOKU-) DOKURITSU GYOSEI HOJIN NOGYO SEIBUTSU SH.
PA (SEIB-) SEIBUTSUKAI TOKUTEI SANGYO GIUTTSU.
PA (DOKU-) DOKURITSU GYOSEI HOJIN RIKAGAKU KENKYUSH.
PA (KOKU-) ZH KOKUSAI KAGAKU SHINKO ZAIDAN.
XX
PI Kikuchi H, Hayashizaki Y, Otsomo Y, Matsumura K, Murekami K,
PI Kishimoto N, Sato K, Nagata T, Kawakami N, Yazaki J, Ishikawa M,
PI Doi K, Kawai J,
XX
DR WPI; 2005-566181/58.

```

PT Novel DNA encoding transcription factor, derived from rice plant, useful
PT for obtaining transcriptional-regulatory regions in plant and for
PT producing modified plant.

XX Claim 1; SEQ ID NO 2746; 2928bp; Japanese.

CC The invention relates to a novel DNA sequence encoding a transcription
CC factor derived from a plant. The invention further comprises antisense
CC RNA sequences, ribozyme activity RNA, RNAi sequences, a vector,
CC transformed plant cells, antibodies and proteins, all related to the
CC novel plant DNA sequences of the invention. The novel DNA is preferably
CC derived from a rice-genome database. The invention further provides a
CC method for determining the transcription regulatory regions of the rice
CC genome. The novel DNA is useful for controlling the expression of a gene
CC in a plant and for producing a modified plant with desired and different
CC characteristics. The plant DNA and method enables the acquisition of many
CC transcriptional-regulatory regions. This polynucleotide represents a DNA
CC sequence taken from a rice genome clone library for use in the invention.
CC Note: This sequence is not shown in the specification. It has been
CC retrieved from a sequence listing in electronic format from the Japanese
CC Patent Office. The invention claims DNA sequences of SEQ ID Nos 1 to
CC 28469 and encoded protein sequences of SEQ ID Nos 28470 to 56791,
CC however, the sequence listing only provided the DNA sequences of SEQ ID
CC Nos 1 to 3032.

XX
SQ Sequence 1137 BP; 244 A; 250 C; 278 G; 365 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,66e-16 Length: 1137
Score: 224.50 Matches: 51
Percent Similarity: 55.7% Conservative: 16
Best Local Similarity: 60.7% Mismatches: 24
Query Match: 42.1% Indels: 11
DB: 14 Gaps: 3

US-10-715-129-2 (1-106) x ABB67601 (1-1137)

QY 3 ValGlyArgAnaAnArgLysAsnMetSerPheArgSerProGlnLysSerIleGln 22
Db 688 GTAGTGTCTAGCAGCAGAGGAAAC---CCTGTGATACCCAGACCAAGCTTCAAGAGG 744
QY 23 GluLeuArgArgSerPhe-----SerIleGlnLysArgMetIleProIleAsn 39
Db 745 CATGTGCAGCATCAGCATCAGCAAAATCTGCTCCAGAGAAAGACGCCGCCAGCATCGAAT 804
QY 40 TyrPheSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuIleLeu 59
Db 805 TACTTCAGTATCGAGGGCTTCCTGCTGCTGCTCTCTCCACACAGTCAATTCCTACTT 864
QY 60 ProLeuValLeuProProLeuProProProPheMetLeuLeuValProIleGly 79
Db 865 CCATTTGGTCTTCCCTCCCTCCGCGCATGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 924
QY 80 IleMetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaThr 99
Db 925 CTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
QY 100 AspVal 101
Db 964 GATGTG 969

RESULT 7
ACL31088/c
ID ACL31088 standard; cDNA; 298 BP.
XX
XX ACL31088;
XX
XX 02-JUN-2005 (first entry)
XX
XX Rice abiotic stress response related polynucleotide SEQ ID NO:9651.
XX
XX ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW agriculture.

XX
OS Oryza sativa.
XX
XX WO2003008540-A2.
XX
XX 30-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-US019668.
XX
XX 22-JUN-2001; 2001US-0300112P.
XX
XX 24-AUG-2001; 2001US-0314662P.
XX
XX 26-SEP-2001; 2001US-0325277P.
XX
XX 21-NOV-2001; 2001US-0332132P.
XX
XX (SYN) SYNGENTA PARTICIPATIONS AG.
XX
XX Krops J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
XX Moughamer T, Provart N, Riecke D, Zhu T;
XX WPI; 2003-248011/24.
XX
XX New stress-responsive nucleic acid, useful for altering the
XX responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
XX stress, salt stress or osmotic stress.
XX
XX Disclosure; SEQ ID NO 9651; 89pp; English.

CC The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention

XX
SQ Sequence 298 BP; 77 A; 71 C; 87 G; 62 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 9.85e-16 Length: 298
Score: 215.00 Matches: 40
Percent Similarity: 83.3% Conservative: 10
Best Local Similarity: 66.7% Mismatches: 10
Query Match: 40.3% Indels: 0
DB: 11 Gaps: 0

US-10-715-129-2 (1-106) x ACL31088 (1-298)

QY 37 ProIleAsnTyrPheSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeu 56
Db 221 CCTCAAAATATCTCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 162
QY 57 LeuIleLeuProLeuValLeuProProLeuProProProPheMetLeuLeuVal 76
Db 161 CTGATCTGCTCTCTCAATCTGCGCAGTCCGCGCGCCGACCTTGATGCTCTGCTGCT 102
QY 77 ProIleGlyIleMetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsn 96
Db 101 CCTATAGGCTCACTTGTGTGCTGATGATCTTCAATGCTGCTGCTGCTGCTGCTGCTGCT 142

RESULT 8
ACL29210
ID ACL29210 standard; cDNA; 978 BP.
XX
XX ACL29210;
XX
XX 02-JUN-2005 (first entry)
XX

XX Plant full length insert polynucleotide seqid 5042.
 DE
 XX

KM plant protectant; plant growth regulator; gene therapy; plant;
 KM recombinant DNA construct; physical array; plant breeding marker;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;
 KM growth rate; cell cycle pathway; disease resistance;
 KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content; gene; ss.

XX Unidentified.
 OS

XX US2004034888-A1.
 PN

XX 19-FEB-2004.
 PD

XX 28-APR-2003; 2003US-00425114.
 PF

XX 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.

XX (LIU/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCREEN/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAO/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 PI WPI; 2004-180133/17.
 XX

XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX

PS Claim 1; SEQ ID NO 5042; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html; DocID:200403488. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 CC

XX Sequence 1157 BP; 225 A; 348 C; 299 G; 285 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 2, 11e-12 Length: 1157
 Score: 193.00 Matches: 44
 Percent Similarity: 60.7% Conservative: 21
 Best Local Similarity: 41.1% Mismatches: 22
 Query Match: 36.2% Indels: 20
 DB: 13 Gaps: 2

US-10-715-129-2 (1-106) x ADX10467 (1-1157)

8 ArglyshamWetSerPheArgSerSerProGluyluSerlysgingluLeuArgArgSer 27

DB 678 AAGAAACATGCGAAGCCGATCTAGCGCATGGAAGAGAGGCGCGCAATCAAGAGAG 737
 QY 28 PhSerAlaGlnlybArgMet----- 34
 DB 738 AATGCCGTGAAAGCATCTCGACGACGCTCAGCAGAGAGCGGATTTCTCGACAAAGAG 797
 QY 35 MetIIProAlaSerlyrPheSerleuGluSerleuPheleuValGlyleuThraAla 54
 DB 798 GTCATGCGCTCCACCTTACATTCAGATCGGAGCGCTTCTCTGCTCGCTCCCTCACCCGTC 857
 QY 55 SerleuLeuIleleuProleuValleuProleuProleuProleuProleuMetleuLeu 74
 DB 858 TCGCTGCTATATCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917
 QY 75 LeuValProIleGlyleuMetValleuLeuValleuAlaPheMetProSerSerHis 94
 DB 918 TGGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 968
 QY 95 SerAsnAlaSerThrAspVal 101
 DB 969 -----ACAGATGTG 977

RESULT 12

ID ADT18995 standard; cDNA, 1055 BP.

XX ADT18995;

DT 13-JAN-2005 (first entry)

XX Plant cDNA, Seq ID 4321.

XX Plant; ss; gene; transgenic; cold tolerance; growth rate;
 KM drought tolerance; disease resistance; galactomannan production;
 KM plant growth regulator; heat tolerance; herbicide tolerance;
 KM lignin production; extreme osmotic condition tolerance;
 KM pathogen resistance; pest resistance; yield improvement; seed oil yield;
 KM seed protein yield.
 XX

OS Viridiplantae.

XX US2004216190-A1.

PD 28-OCT-2004.

PF 18-DEC-2003; 2003US-00739930.

XX 28-APR-2003; 2003US-00424599.

PR 28-APR-2003; 2003US-00425115.

XX (KOVA/) KOVALIC D K.

XX Kovalic DK;

PI WPI; 2004-757369/74.

XX New recombinant DNA constructs useful in the field of biochemistry and
 PT genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.
 XX

PS Claim 1; SEQ ID NO 4321; 14pp; English.

XX The invention relates a recombinant DNA construct comprising a
 CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
 CC Arabidopsis, wheat and rape but the specification does not indicate which
 CC sequences is derived from which organism. Also included is a method of
 CC producing a plant having an improved property, comprising transforming a
 CC plant with a recombinant DNA construct comprising a promoter region
 CC functional in a plant cell operably joined to a polynucleotide encoding a
 CC polypeptide associated with the property, and growing the transformed

CC plant. The property is selected from improving plant cold tolerance, for
 CC manipulating growth rate in plant cells by modification of the cell cycle
 CC pathway, for improving plant drought tolerance, for providing increased
 CC resistance to plant disease, for galactomannan production, for production
 CC of plant growth regulators, for improving plant heat tolerance, for
 CC improving plant tolerance to herbicides, for increasing the rate of
 CC homologous recombination in plants, for lignin production, for improving
 CC plant tolerance to extreme osmotic conditions, for improving plant
 CC tolerance to pathogens or pests, for yield improvement by modification of
 CC photosynthesis, for modifying seed oil yield and/or content, for
 CC modifying seed protein yield and/or content, for yield improvement by
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
 CC and for yield improvement by providing improved plant growth and
 CC development under at least one stress condition. The polynucleotide may
 CC also encode a plant transcription factor. The methods and compositions of
 CC the present invention are useful in the field of biochemistry and
 CC genetics, in particular for producing transgenic plants with improved
 CC biological characteristics such as increased yield, improved nitrogen
 CC flow, increasing plant tolerance to cold or heat, improving plant
 CC tolerance to extreme osmotic and drought conditions, and improving plant
 CC tolerance to plant pests or pathogens. They can also be used in physical
 CC arrays of molecules, plant breeding markers, computer-based storage and
 CC analysis systems. The present sequence is one of the 5544 plant cDNA
 CC sequences of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from USPTO at
 CC Seqdata.uspto.gov/sequence.html?docID=20040216190.

XX Sequence 1055 BP; 218 A; 331 C; 268 G; 238 T; 0 U; 0 Other;

Alignment Scores:

Score: 4.57e-11 Length: 1055
 181.00 Matches: 45
 Percent Similarity: 63.1% Conservative: 20
 Best Local Similarity: 43.7% Mismatches: 33
 Query Match: 34.0% Indels: 5
 DB: 13 Gaps: 2

US-10-715-129-2 (1-106) x ADT18995 (1-1055)

QY 5 ArgAnbAnArgLyAsnMet-SerPheArgSerProGluLySerLySGInGluLe 24
 Db 526 AGGAGCGCGCGGAGACAGCTGCTGCGCATCGCCGCGTGGTCCGCGCGGACCA 585
 QY 24 uArgARSerPheSerAlaGlnLyS-----ArgMetMetLeProAlaenTYrph 41
 Db 586 GAAAGCAACGATACAGGAGGAGGAGACCCGCGTGTGACAGCTGACGAGTACGCTCTT 645
 QY 41 e---SerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuLeuPr 60
 Db 646 CGGCAAGTACTTCTCTGCTCTCTGCTCATATGTCACCGCGTGTGATTTCTCC 705
 QY 60 OlauValLeuProProLeuProProProPheMetLeuLeuValProIleGlyIi 80
 Db 706 GCTGTCTCTGACACCGCTGCTCCGCCCGCTGCATGCTATGCTGGTCCAGTGGCAAT 765
 QY 80 emetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAlaAlaenThrAs 100
 Db 766 GCTGTGATGCTCTGCTGCTGCGTTCATCCGACGTGAGCGCGGGGCGACGCGG 825
 QY 100 pValThr 102
 Db 826 CCCGACT 832

RESULT 13
 ADT17295/c
 ID ADT17295 standard; cDNA; 939 BP.

XX AC ADT17295;
 XX DT 13-JAN-2005 (first entry)
 XX DB Plant cDNA, Seq ID 2621.

XX Plant; ser; gene; transgenic; cold tolerance; growth rate;
 KW drought tolerance; disease resistance; galactomannan production;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW lignin production; extreme osmotic condition tolerance;
 KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
 KW seed protein yield.
 OS Vitidiplantae.
 PN US2004216190-A1.
 XX
 PD 28-OCT-2004.
 XX
 PP 18-DEC-2003; 2003US-00739930.
 XX
 PR 28-APR-2003; 2003US-00424599.
 XX
 PR 28-APR-2003; 2003US-00425115.
 XX
 PA (KOVA/) KOVALIC D K.
 XX
 X1 Kovalic DK;
 XX
 DR WPI; 2004-757369/74.
 XX
 PT New recombinant DNA constructs useful in the field of biochemistry and
 PT genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.
 XX
 PS Claim 1; SEQ ID NO 2621; 14pp; English.

XX The invention relates a recombinant DNA construct comprising a
 CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
 CC Arabidopsis, wheat and rape but the specification does not indicate which
 CC sequences is derived from which organism. Also included is a method of
 CC producing a plant having an improved property, comprising transforming a
 CC plant with a recombinant DNA construct comprising a promoter region
 CC functional in a plant cell operably joined to a polynucleotide encoding a
 CC polypeptide associated with the property, and growing the transformed
 CC plant. The property is selected from improving plant cold tolerance, for
 CC manipulating growth rate in plant cells by modification of the cell cycle
 CC pathway, for improving plant drought tolerance, for providing increased
 CC resistance to plant disease, for galactomannan production, for production
 CC of plant growth regulators, for improving plant heat tolerance, for
 CC improving plant tolerance to herbicides, for increasing the rate of
 CC homologous recombination in plants, for lignin production, for improving
 CC plant tolerance to extreme osmotic conditions, for improving plant
 CC tolerance to pathogens or pests, for yield improvement by modification of
 CC photosynthesis, for modifying seed oil yield and/or content, for
 CC modifying seed protein yield and/or content, for yield improvement by
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
 CC and for yield improvement by providing improved plant growth and
 CC development under at least one stress condition. The polynucleotide may
 CC also encode a plant transcription factor. The methods and compositions of
 CC the present invention are useful in the field of biochemistry and
 CC genetics, in particular for producing transgenic plants with improved
 CC biological characteristics such as increased yield, improved nitrogen
 CC flow, increasing plant tolerance to cold or heat, improving plant
 CC tolerance to extreme osmotic and drought conditions, and improving plant
 CC tolerance to plant pests or pathogens. They can also be used in physical
 CC arrays of molecules, plant breeding markers, computer-based storage and
 CC analysis systems. The present sequence is one of the 5544 plant cDNA
 CC sequences of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from USPTO at
 CC Seqdata.uspto.gov/sequence.html?docID=20040216190.

XX Sequence 939 BP; 215 A; 263 C; 287 G; 174 T; 0 U; 0 Other;

Alignment Scores: 2.43e-09 Length: 939
 Pred. No.: 2.43e-09

Score: 165.50 Matches: 44
 Percent Similarity: 60.6% Conservative: 19
 Best Local Similarity: 42.3% Mismatches: 38
 Query Match: 31.1% Indels: 4
 DB: 13 Gaps: 2

US-10-715-129-2 (1-106) x ADT17295 (1-939)

Qy 5 ArgAsnAsnArgLysAsnMetSerPhe-----ArgSerSerProGluLysSerLysGln 22
 Db 433 CAGAAATGAGAGCGACCATGCTGCTGCAAGAAAGAGCCCGACGGTGGCCAAAGGG 374
 Qy 23 GluLeuArgArgSerPheSerAlaGlnLysArgMetMetIleProAlaAsnTyrPheSer 42
 Db 373 CAGAGCGGTGAGAAAGCTTGGCCCGCTGCGCCCAAGCTCGCCCGGGGGGCTCAGC 314
 Qy 43 LeuGlnSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuIleLeuProLeuVal 62
 Db 313 GCGAGAGCGTTCCTGCTTCTGCGTGGCGCCGCTGCTTATCGTGGTCCGCTGCTC 254
 Qy 63 LeuProProLeuProProProPheMetLeuLeuValProIleGlyIleMetVal 82
 Db 253 CTGCGCGCGCTGCGCGCCCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194
 Qy 83 LeuLeuValValLeuAla---PheMetProSerSerHisSerAsnAlaAsnThrAspVal 101
 Db 193 CTCCTCGCGCGCTGCGCCACCTTCATGCGCTGAGTGC--TCAGAGCATGCGATCTCCA 136
 Qy 102 ThrCysAsnPhe 105
 Db 135 ACTGTAACTAC 124

RESULT 14
 ADX37027/c
 ID ADX37027 standard; cDNA; 936 BP.

AC ADX37027;

DT 21-APR-2005 (first entry)

XX plant full length insert polynucleotide seqid 19847.

XX plant protectant; plant growth regulator; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/J) LIU J.

XX (ZHOV/J) ZHOV Y.

XX (KOVA/J) KOVALIC D K.

XX (SCRE/J) SCREEN S E.

XX (TABY/J) TABASKA J E.

XX (CAOV/J) CAO Y.

XX LIU J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.

XX Claim 1; SEQ ID NO 19847; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.seqdata.uspo.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 936 BP; 247 A; 221 C; 234 G; 234 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.:	4.55e-08	Length:	936
Score:	154.50	Matches:	33
Percent Similarity:	60.0%	Conservative:	9
Best Local Similarity:	47.1%	Mismatches:	17
Query Match:	29.0%	Indels:	11
DB:	13	Gaps:	1

US-10-715-129-2 (1-106) x ADX37027 (1-936)

Qy 25 ArgArgSerPheSerHisGlnLysArgMetMetIleProAlaAsnTyrPheSerLeuGlu 44

Db 746 GCGAGGCGCGCGCGATGAGCAAGAGATG----- 717

Qy 45 SerLeuPheLeuValGlyLeuThrAlaSerLeuLeuIleLeuProLeuValLeuPro 64

Db 716 ---CTATGATGATGCTGCGCGGAGACAGTATCTCTGCTGCTGCTGCTGCTGCTGCA 660

Qy 65 ProLeuProProProPheMetLeuLeuValProIleGlyIleMetValLeuLeu 84

Db 659 CCCCTCCCGCACCGCGCTGTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 600

Qy 85 ValValLeuAlaPheMetProSerSerHis 94

Db 599 TTCTCCTGCTGTTTCTTCTCCGCTCAACAC 570

XX RESULT 15

XX ADR64525 standard; cDNA; 366 BP.

XX ADR64525;

XX 02-DEC-2004 (first entry)

XX Cotton cDNA sequence, SEQ ID 5306.

XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
 KW drought tolerance; plant disease resistance; galactomannan; lignin;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW homologous recombination; extreme osmotic condition tolerance;
 KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
 KW stress resistance.

XX Gossypium hirsutum.

XX OS

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: January 30, 2006, 09:13:58 ; Search time 148 Seconds
(without alignments)
1273.118 Million cell updates/sec

Title: US-10-715-129-2

Perfect score: 533

Sequence: 1 MDVGRNRRKXKMSFRSPSPKS.....LAFMPSHSNANDVTGCMF 106

Scoring table: BIOSUM62
Xgapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=x1p
-Q=/cgn2_1/USPRO.spool/p/US10715129/runat_27012006_154745_14592/abp_query.fasta_1.263
-DB=Issued_Patents_NA -OPMT=fastcap -SUPPLX=p2n.rml -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US10715129.@CGN_1_1_230@runat_27012006_154745_14592 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:
1: /cgn2_6/ptodata/1/ina/1.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.5	17.9	312957	3	US-09-949-001-31
2	95.5	17.9	312972	3	US-09-949-001-34
3	95	17.8	5340	3	US-09-535-145-1
4	95	17.8	44166	3	US-09-949-016-15829
5	95	17.8	54180	3	US-09-949-016-14894
6	93.5	17.5	152582	3	US-09-949-016-12086
7	93.5	17.5	152583	3	US-09-949-016-17390
8	93.5	17.5	152583	3	US-09-949-016-17391
9	92	17.3	601	3	US-09-949-016-115267

c 10	92	17.3	601	3	US-09-949-016-144537	Sequence 144537, A
c 11	92	17.3	165651	3	US-09-949-016-13032	Sequence 11032, A
c 12	90	16.9	601	3	US-09-949-016-180704	Sequence 180704, A
c 13	90	16.9	24979	2	US-08-147-777-3	Sequence 3, Appl1
c 14	90	16.9	24979	3	US-08-452-872-3	Sequence 3, Appl1
c 15	90	16.9	24979	6	PCT-US93-03985-3	Sequence 3, Appl1
c 16	90	16.9	70770	3	US-09-949-016-16938	Sequence 16938, A
c 17	89.5	16.8	3870	3	US-09-949-016-5555	Sequence 5555, Ap
c 18	89.5	16.8	3901	3	US-09-949-016-448	Sequence 448, App
c 19	89.5	16.8	16885	3	US-09-949-016-17297	Sequence 17297, A
c 20	89.5	16.8	16880	3	US-09-949-016-12190	Sequence 12190, A
c 21	87.5	16.4	601	3	US-09-949-016-16420	Sequence 16420, A
c 22	86.5	16.2	234884	3	US-09-949-016-16420	Sequence 16420, A
c 23	85.5	16.0	18798	3	US-09-949-016-14339	Sequence 14339, A
c 24	85.5	16.0	85850	3	US-09-949-016-13424	Sequence 13424, A
c 25	85	15.9	247781	3	US-09-949-016-14193	Sequence 14193, A
c 26	85	15.9	767677	3	US-09-949-016-12147	Sequence 12147, A
c 27	85	15.9	767677	3	US-09-949-016-17361	Sequence 17361, A
c 28	84.5	15.9	2093	3	US-09-620-312D-999	Sequence 999, App
c 29	84.5	15.9	2358	3	US-09-248-786A-6328	Sequence 6328, App
c 30	84.5	15.9	18955	3	US-09-949-016-13343	Sequence 13343, A
c 31	84.5	15.9	30678	3	US-09-949-016-12818	Sequence 12818, A
c 32	84	15.8	15252	3	US-09-949-016-13584	Sequence 13584, A
c 33	84	15.8	387902	3	US-09-949-016-14543	Sequence 14543, A
c 34	84	15.8	421883	3	US-09-949-016-12557	Sequence 12557, A
c 35	83.5	15.7	601	3	US-09-949-016-117038	Sequence 117038, A
c 36	83.5	15.7	1012	3	US-09-775-398-56	Sequence 66, Appl
c 37	83.5	15.7	3466	3	US-10-104-047-536	Sequence 536, App
c 38	83.5	15.7	3494	3	US-09-949-016-4323	Sequence 4323, Ap
c 39	83.5	15.7	5316	3	US-09-814-915A-26	Sequence 26, Appl
c 40	83.5	15.7	194937	3	US-09-949-016-17032	Sequence 17032, A
c 41	83.5	15.7	194937	3	US-09-949-016-17033	Sequence 17033, A
c 42	83.5	15.7	212139	3	US-09-949-016-16065	Sequence 16065, A
c 43	83.5	15.7	325034	3	US-09-949-016-14957	Sequence 14957, A
c 44	83.5	15.7	389504	3	US-09-949-016-11774	Sequence 11774, A
c 45	83	15.6	2554	3	US-09-799-451-716	Sequence 716, App

ALIGNMENTS

RESULT 1
US-09-949-001-31
; Sequence 31, Application US/09949001
; Patent No. 6625336
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1000789
; CURRENT APPLICATION NUMBER: US/09/949, 001
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 312957
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(312957)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-001-31

Alignment Scores:
Pred. No.: 69.5
Score: 95.50
Percent Similarity: 57.1%
Best Local Similarity: 42.9%
Query Match: 17.9%
DB: 3
Length: 312957
Matches: 27
Conservative: 9
Mismatch: 25
Indels: 2
Gaps: 1

[illegible]

SEQ ID NO 15829
LENGTH: 44166
TYPE: DNA
ORGANISM: Human
US-09-949-016-15829

Alignment Scores:
Pred. No.: 4.51 Length: 44166
Score: 95.00 Matches: 29
Percent Similarity: 51.8% Conservative: 15
Best Local Similarity: 34.1% Mismatches: 33
Query Match: 17.8% Indels: 8
DB: 3 Gaps: 3

US-10-715-129-2 (1-106) x US-09-949-016-15829 (1-44166)

Qy 14 ArgSerSerProGluYSerLysGlnGluLeuArgSerPheSerAlaGlnLysArg 33
Db 4671 AGGTTTCTCTCTGACACACACCTCCCAAGTAGCTGACTTTT-----AAAAATGT 4724
Qy 34 MetMetIleProAlaSerTyPheSerLeuGluSerLeuPheLeu-----Val 50
Db 4725 TTCTGTGTTCTGCTGTTCTTCTTCTATGTTGTTCTTCTCTCTCTCTCT 4784
Qy 51 GlyLeuThrAlaSerLeuLeuLeuProLeuValLeuProProLeuProProPro 70
Db 4785 GTTGTGTATTTCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4844
Qy 71 PheMetLeuLeuValProIleGlyIleMetValLeuValLeuAlaPheMet 90
Db 4845 CT 4895
Qy 91 ProSerSerHisSer 95
Db 4896 TCCCT 4910

RESULT 5
US-09-949-016-14894

Sequence 14894, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14894
LENGTH: 54180
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(54180)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14894

Alignment Scores:
Pred. No.: 6.09 Length: 54180
Score: 95.00 Matches: 29
Percent Similarity: 51.8% Conservative: 15
Best Local Similarity: 34.1% Mismatches: 33
Query Match: 17.8% Indels: 8
DB: 3 Gaps: 3

US-10-715-129-2 (1-106) x US-09-949-016-14894 (1-54180)

Qy 14 ArgSerSerProGluYSerLysGlnGluLeuArgSerPheSerAlaGlnLysArg 33
Db 14825 AGGTTTCTCTCTGACACACCTCCCAAGTAGCTGACTTTT-----AAAAATGT 14878
Qy 34 MetMetIleProAlaSerTyPheSerLeuGluSerLeuPheLeu-----Val 50
Db 14879 TTCTGTGTTCTGCTGTTCTTCTTCTATGTTGTTCTTCTCTCTCTCTCTCTCT 14938
Qy 51 GlyLeuThrAlaSerLeuLeuLeuProLeuValLeuProProLeuProProPro 70
Db 14939 GTTGTGTATTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 14998
Qy 71 PheMetLeuLeuValProIleGlyIleMetValLeuValLeuAlaPheMet 90
Db 14999 CT 15049
Qy 91 ProSerSerHisSer 95
Db 15050 TCCCT 15064

RESULT 6
US-09-949-016-12086/c

Sequence 12086, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12086
LENGTH: 152582
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(152582)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12086

Alignment Scores:
Pred. No.: 43.5 Length: 152582
Score: 93.50 Matches: 37
Percent Similarity: 52.0% Conservative: 15
Best Local Similarity: 37.0% Mismatches: 29
Query Match: 17.5% Indels: 19
DB: 3 Gaps: 5

US-10-715-129-2 (1-106) x US-09-949-016-12086 (1-152582)

Qy 3 ValGlyArgAlaSerLeuValGlnMetSerPheArgSerProGluYSerLysGln 22
Db 108711 GTAAATGCCACACCTCAAGAGTCACACCTGAGGCTCTCCCAATTAAGCT---ACT 108655
Qy 23 GluLeuArgSerPheSerAlaGlnLysArgMetIleProAlaSerTyPheSer 42
Db 108654 TCCATGTTTGAAGATGAGGCTCTCTCAAGAGA-----AACTTCTTCTCC 108610
Qy 43 LeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeu-----Le 57
Db 108609 TTCTTCT 108556


```

; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-115267

```

Alignment Scores:	
Pred. No.:	0.0195
Score:	92.00
Percent Similarity:	51.8%
Best local Similarity:	33.7%
Query Match:	17.3%
DB:	3
Length:	603
Matches:	15
Conservative:	28
Mismatches:	32
Indels:	
Gaps:	3

US-10-715-129-2 (1-106) X US-09-949-016-115267 (1-601)

Oy 14 ArgSerSerProGluIlySerIySgIngluLeuArgArgSerPheSerIaGInIlybAg 33
 ||| ||||| :: |||
 Db 236 AGGTTTTCCTCAGACACCACTCCCAAGTAGCTGACTTTT-----AAAAATGT 189

```
Oy      34 MetcetiIleproAlabnTYrPheSerLeugInUSeuriePhelu-----Val 50
          :: ||| :||:::||::||| |::||| ::|
Db     182 TTGTGTTCTGCCTGGTTCATTATCATCAATGGTGTTCTCCTCCCTCTCT 122
```

Qy	51 GlyLeuThrIaSerLeuLeuLeuProleuValLeuProProleuProProPro	70
	:::	
Db	122 GTGTGTTATCTTCTTTCTTCCTCCTCCTCCTCCTCCTCCTCCTCCT	63

```

Qy      71 PheMetLeuLeuValProIleGlyIleMetValLeuLeuValIleLeuAlaPheMet 90
      :: ||| ::::: ||| ::|
Db      62 CTTCCTCTCTCTCTTC-----CTTCCTCTCTCTCTTCACCTTCACCTTCGCC 12

```

Qy	91 ProSerSer 93
Db	11 TCCTCTCC 3

RESULT 10
US-09-949-016-144537/c
; Sequence 144537, Application US/09949016
; Patent No. 681330

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY FILING DATE: 2000-04-14

```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 144537
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; OS-09-949-016-144537

```

US-10-715-129-2 (1-106) . X US-09-949-016-144537 (1-601)

```

Oy      14 ArgSerSerProGluIuIysSerIysGInGluLeuArgSerPheSerAlaGlnIuIysArg 33
          ||| |||||
Db      236 AGGTTTCTCCTCATGACACCACTCCCAAGTAGCTGCACTTTT-----AAAAATCT 18
          ||| |||||

```

```
Oy      34 MetMetIleProAlaAsnTrpPheSerLeuGlnSerLeuPheLeu-----Val 50
          ::| |::|||::| | ||||| | |
Db      182 TTTGTTCCTCAGCTGTTCTTTCATCATATGGTGTTTCTTCTCCCTCTT 12
```

```

Qy      51 GlyLeuThrAlaSerLeuLeuLeuProLeuValLeuProProLeuProProPro 70
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      122 GTGTGTGTAATCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 63

```

Oy		71	PheMetLeuValProIleGlyIleMetValLeuValValLeuAlaPheMet	90
	:::		::::: :::	
Dd		62	CYTCTCTTCTTC-----CTTCCTCCTCCTCCACACTTCACCTTCGCC	12

Qy	91 ProSerSer 93
Db	11 TCCTCTCTCC 3

RESULT 11
US-09-949-016-13032
; Sequence 13032, Application US/09949016

```

: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14

```

```

? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSQ for Windows Version 4.0.
? SEQ ID NO 13032
? LENGTH: 165651
? TYPE: DNA
? ORGANISM: Human
? FEATURE:

```

```

; NAME/NEI: mibc_readuc
; LOCATION: (1)...(165651)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13032

```

Alignment Scores:	
Prod. No.:	76.4
Score:	92.00
Percent Similarity:	49.58
Best Local Similarity:	34.08
Query Match:	17.38
DB:	
US-10-715-129-2 (1-106)	
US-09-949-016-13032 (1-165551)	
Length:	11
Matches:	3
Conservative:	1
Mismatches:	2
Indels:	2
Gaps:	5

	QY	15 SerSerProGluIuysSerLysGlnGluleArgArsgSerPheSeraAlagInllyeargMet	34
	D8	116094 TCTTCTCCTTCCTTTCTTCCTCCCTCCCTCAACTTCTCTCT	116141 :::

```
QY      35 MetLeIProLalaAntyRPhseSerLeugUserleupheutLeuValGlyLeuThrAla   54
          ::::| | | | | | | | | | | | | | | | | | | | | | | | | |
Db      116142 -----TCGCTTCTTCTCTCTTTCTCTCTCTCTCCCCCTCGCTCTC----- 116180
```

```
QY      55 SerLeuIleLeuProLeuValLeuProProLeuProProProPhe-----Met 72
          :::::
Dbb    116181 ---CCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCATCATCGT 116231
```

Oy		73	LleuleuValProIleGlyIleMetValLleuleuValVal-----86
		:	: : : : : :
Ddb		116232	CCTTCCTTCCCTTGAGCCTTAGTTACTTCCTCGCTTCCTCCTCACTCCC116291

QY 87 -----LeuAlaPheMetProSerSerHisSerAsnAlaAsnThrAspVal 101

```

Db      116292 CAAACCCTCCCTCCTCCTCCTCCCTCCACATCGTCAACTCTTTCTTCCTGT 116342
RESULT 12
US-09-949-016-180704
Sequence 180704, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: GL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 180704
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-180704

Alignment Scores:
Pred. No.:          0.0352           Length:        601
Score:             90.00            Matches:       25
Percent Similarity: 47.6%           Conservative:   15
Best Local Similarity: 29.8%         Mismatches:    18
Query Match:       16.9%           Indels:        26
DB:                3               Gaps:          2
US-10-715-129-2 (1-106) x US-09-949-016-180704 (1-601)

Qy      MetMetIleProAlaAsnTyRPheserLeugLUseriePheuleValGlyeauthr 53
Dy      CTACGCCTGCCACTACTATAAAGTACTACCATTAAGAATTCTACTACTATATACATACTACTACTA 105
Qy      A1asesterleulle-----leuproleuvalleupropro----- 65
Dy      TTGTCTACTACTACTACTACTACTACATTAAGTAGTACACACACTATCATCATCATCATCACTACTA 165
Qy      ----- 66 ----- -leupro 67
Dy      CTAAATACTATCATTAACACTACTACTATTAACAATCACAGACTATCATTAATCACTACTACTACCA 225
Qy      ProProProPheMetleulleuleuValProIIeglylleMetValleuleuValleu 87
Dy      CCTCACCATTAATCTACTACTATTAATGCTATCATTAATCACTACTCTGGTCTCTATATTAATCACTAC 285
Qy      88 AlphaMetPro 91
Dy      286 CTACTCTCACTCA 297

RESULT 13
US-08-147-777-3
Sequence 3, Application US/08147777
Patent No. 5914265
GENERAL INFORMATION:
APPLICANT: Roop, Dennis R.
APPLICANT: Rothmagerl, Joseph A.
APPLICANT: Greenhalgh, David A.
APPLICANT: Yuspa, Stuart H.
TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: LYON & LYON
STREET: 611 West Sixth Street

```

```

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/147,777
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 07/876,289
FILING DATE: April 30, 1992
APPLICATION NUMBER: Unassigned (204/144)
FILING DATE: October 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,337
REFERENCE/DOCKET NUMBER: 204/153
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 24979 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-147-777-3
Alignment Scores:
Pred. No.: 8.5 Length: 24979
Score: 90.00 Matches: 26
Percent Similarity: 53.9% Conservative: 15
Best local Similarity: 34.2% Mismatches: 9
Query Match: 16.9% Indels: 26
DB: Gaps: 3
US-10-715-129-2 (1-106) x US-08-147-777-3 (1-24979)
Cy 41 PhaeSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuPro 60
Db 6874 TGGCAGAGATGATATTTCTTCTT-----CTTCTCTTCTCTC 6918
Cy 61 LeuValLeuProPheLeuProProProPheMetLeuLeuValProIleGly--- 79
Db 6919 CTTTGCTTCTCTCTCTCTTCCCTCC-----CTGCTTCTCTTGCTCTCGGCTCT 6972
Cy 80 -----IleMet 81
Db 6973 CTTCCGCGCTCTCTCTCTCTCTCTCCGCGCTCTCTCCGCGCGCTCTCTCTCTT 7032
Cy 82 ValLeuLeuValValLeuAlaPheMetProSerSerHisSerAnaIa 97
Db 7033 CTTCTTCTCTCTCTCTCTCTTCATCTTCCATCTTCCATCTTCACTTCTCTCT 7080
RESULT 14
US-08-452-872-3
Sequence 3, Application US/08452872
Patent No. 6057298
GENERAL INFORMATION:
APPLICANT: Roop, Dennis R.
APPLICANT: Rothnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
APPLICANT: Yuepa, Stuart H.
TITLE OF INVENTION: KERATIN KI EXPRESSION VECTORS

```

US-10-715-129-2 (1-106) X US-08-452-872-3 (1-24979)

RESULT15
PCT-US93-03985-3
; Sequence 3, Application PC/TUS9303985
; GENERAL INFORMATION:
; APPLICANT: Roop, Dennis R.

US-10-715-129-2 (1-106) X PCT-US93-03985-3 (1-24979)

Search completed: January 30, 2006, 11:22:39
Job time : 300 secs

THIS PAGE BLANK (uspto)